

ORGANIZADORES

Fabiano Couto Corrêa da Silva

Thiago Henrique Bragato Barros

CAMINHOS METODOLÓGICOS EM CIÊNCIA DA INFORMAÇÃO

fundamentos
e aplicações



CAPES



PPGCIN
Programa de Pós-Graduação
em Ciência da Informação



pimenta
cultural

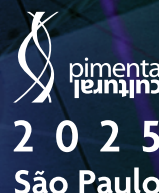
ORGANIZADORES

Fabiano Couto Corrêa da Silva

Thiago Henrique Bragato Barros

CAMINHOS METODOLÓGICOS EM CIÊNCIA DA INFORMAÇÃO

fundamentos
e aplicações



DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

C183

Caminhos Metodológicos em Ciência da Informação:
fundamentos e aplicações / Organização Fabiano Couto
Corrêa da Silva, Thiago Henrique Bragato Barros. – São
Paulo: Pimenta Cultural, 2025.

Livro em PDF

ISBN 978-85-7221-475-9

DOI 10.31560/pimentacultural/978-85-7221-475-9

1. Metodologia de pesquisa. 2. Ciência da Informação.
3. Estudos métricos. 4. Ciência Aberta. 5. Preservação digital.
I. Silva, Fabiano Couto Corrêa da (Org.). II. Barros, Thiago
Henrique Bragato (Org.). III. Título.

CDD 025.84

Índice para catálogo sistemático:

I. Ciência da Informação

II. Preservação digital

Simone Sales - Bibliotecária - CRB ES-000814/0

Copyright © Pimenta Cultural, alguns direitos reservados.

Copyright do texto © 2025 os autores e as autoras.

Copyright da edição © 2025 Pimenta Cultural.

Esta obra é licenciada por uma Licença Creative Commons:

Atribuição-NãoComercial-SemDerivações 4.0 Internacional - (CC BY-NC-ND 4.0).

Os termos desta licença estão disponíveis em:

<<https://creativecommons.org/licenses/>>.

Direitos para esta edição cedidos à Pimenta Cultural.

O conteúdo publicado não representa a posição oficial da Pimenta Cultural.

Direção editorial	Patrícia Bieging Raul Inácio Busarello
Editora executiva	Patrícia Bieging
Gerente editorial	Landressa Rita Schiefelbein
Assistente editorial	Júlia Marra Torres
Estagiária editorial	Ana Flávia Pivisan Kobata
Diretor de criação	Raul Inácio Busarello
Assistente de arte	Naiara Von Groll
Editoração eletrônica	Andressa Karina Voltolini
Estagiária em editoração	Stela Tiemi Hashimoto Kanada
Imagens da capa	raftel - Freepik.com
Tipografias	Acumin, Agenda One, Trump Soft Pro
Revisão	Os organizadores
Organizadores	Fabiano Couto Corrêa da Silva Thiago Henrique Bragato Barros

PIMENTA CULTURAL
São Paulo • SP
+55 (11) 96766 2200
livro@pimentacultural.com
www.pimentacultural.com



CONSELHO EDITORIAL CIENTÍFICO

Doutores e Doutoradas

Adilson Cristiano Habowski <i>Universidade La Salle, Brasil</i>	Bernadette Beber <i>Universidade Federal de Santa Catarina, Brasil</i>
Adriana Flávia Neu <i>Universidade Federal de Santa Maria, Brasil</i>	Bruna Carolina de Lima Siqueira dos Santos <i>Universidade do Vale do Itajaí, Brasil</i>
Adriana Regina Vettorazzi Schmitt <i>Instituto Federal de Santa Catarina, Brasil</i>	Bruno Rafael Silva Nogueira Barbosa <i>Universidade Federal da Paraíba, Brasil</i>
Aguimario Pimentel Silva <i>Instituto Federal de Alagoas, Brasil</i>	Caio Cesar Portella Santos <i>Instituto Municipal de Ensino Superior de São Manuel, Brasil</i>
Alaim Passos Bispo <i>Universidade Federal do Rio Grande do Norte, Brasil</i>	Carla Wanessa do Amaral Caffagni <i>Universidade de São Paulo, Brasil</i>
Alaim Souza Neto <i>Universidade Federal de Santa Catarina, Brasil</i>	Carlos Adriano Martins <i>Universidade Cruzeiro do Sul, Brasil</i>
Alessandra Knoll <i>Universidade Federal de Santa Catarina, Brasil</i>	Carlos Jordan Lapa Alves <i>Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil</i>
Alessandra Regina Müller Germani <i>Universidade Federal de Santa Maria, Brasil</i>	Caroline Chioquetta Lorenset <i>Universidade Federal de Santa Catarina, Brasil</i>
Aline Corso <i>Universidade do Vale do Rio dos Sinos, Brasil</i>	Cassia Cordeiro Furtado <i>Universidade Federal do Maranhão, Brasil</i>
Aline Wendpap Nunes de Siqueira <i>Universidade Federal de Mato Grosso, Brasil</i>	Cássio Michel dos Santos Camargo <i>Universidade Federal do Rio Grande do Sul, Brasil</i>
Ana Rosangela Colares Lavand <i>Universidade Estadual do Norte do Paraná, Brasil</i>	Cecilia Machado Henriques <i>Universidade Federal de Santa Catarina, Brasil</i>
André Gobbo <i>Universidade Federal da Paraíba, Brasil</i>	Christiano Martino Otero Avila <i>Universidade Federal de Pelotas, Brasil</i>
André Tanus Cesário de Souza <i>Faculdade Anhanguera, Brasil</i>	Cláudia Samuel Kessler <i>Universidade Federal do Rio Grande do Sul, Brasil</i>
Andressa Antunes <i>Universidade Federal de Ouro Preto, Brasil</i>	Cristiana Barcelos da Silva <i>Universidade do Estado de Minas Gerais, Brasil</i>
Andressa Wiebusch <i>Universidade Federal de Santa Maria, Brasil</i>	Cristiane Silva Fontes <i>Universidade Federal de Minas Gerais, Brasil</i>
Andreza Regina Lopes da Silva <i>Universidade Federal de Santa Catarina, Brasil</i>	Daniela Susana Segre Guertzenstein <i>Universidade de São Paulo, Brasil</i>
Angela Maria Farah <i>Universidade de São Paulo, Brasil</i>	Daniele Cristine Rodrigues <i>Universidade de São Paulo, Brasil</i>
Anísio Batista Pereira <i>Universidade do Estado do Amapá, Brasil</i>	Dayse Centurion da Silva <i>Universidade Anhanguera, Brasil</i>
Antonio Edson Alves da Silva <i>Universidade Estadual do Ceará, Brasil</i>	Dayse Sampaio Lopes Borges <i>Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil</i>
Antonio Henrique Coutelo de Moraes <i>Universidade Federal de Rondonópolis, Brasil</i>	Deilson do Carmo Trindade <i>Instituto Federal de Educação, Ciência e Tecnologia do Amazonas, Brasil</i>
Arthur Vianna Ferreira <i>Universidade do Estado do Rio de Janeiro, Brasil</i>	Diego Pizarro <i>Instituto Federal de Brasília, Brasil</i>
Ary Albuquerque Cavalcanti Junior <i>Universidade Federal de Mato Grosso, Brasil</i>	Dorama de Miranda Carvalho <i>Escola Superior de Propaganda e Marketing, Brasil</i>
Asterlindo Bandeira de Oliveira Júnior <i>Universidade Federal da Bahia, Brasil</i>	Edilson de Araújo dos Santos <i>Universidade de São Paulo, Brasil</i>
Bárbara Amaral da Silva <i>Universidade Federal de Minas Gerais, Brasil</i>	Edson da Silva <i>Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brasil</i>

Elena Maria Mallmann

Universidade Federal de Santa Maria, Brasil

Eleonora das Neves Simões

Universidade Federal do Rio Grande do Sul, Brasil

Eliane Silva Souza

Universidade do Estado da Bahia, Brasil

Elvira Rodrigues de Santana

Universidade Federal da Bahia, Brasil

Estevão Schultz Campos

Centro Universitário Adventista de São Paulo, Brasil

Éverly Pegoraro

Universidade Federal do Rio de Janeiro, Brasil

Fábio Santos de Andrade

Universidade Federal de Mato Grosso, Brasil

Fabrícia Lopes Pinheiro

Universidade Federal do Estado do Rio de Janeiro, Brasil

Fauston Negreiros

Faustidade de Brasília, Brasil

Felipe Henrique Monteiro Oliveira

Universidade Federal da Bahia, Brasil

Fernando Vieira da Cruz

Universidade Estadual de Campinas, Brasil

Flávia Fernanda Santos Silva

Universidade Federal do Amazonas, Brasil

Gabriela Moysés Pereira

Universidade Federal do Rio de Janeiro, Brasil

Gabriella Eldereti Machado

Universidade Federal de Santa Maria, Brasil

Germano Ehler Pollnow

Universidade Federal de Pelotas, Brasil

Geuciane Felipe Guerim Fernandes

Universidade Federal do Pará, Brasil

Geymeesson Brito da Silva

Universidade Federal de Pernambuco, Brasil

Giovanna Ofretorio de Oliveira Martin Franchi

Universidade Federal de Santa Catarina, Brasil

Handherson Leylton Costa Damasceno

Universidade Federal da Bahia, Brasil

Hebert Elias Lobo Sosa

Universidad de Los Andes, Venezuela

Helciclever Barros da Silva Sales

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasil

Helena Azevedo Paulo de Almeida

Universidade Federal de Ouro Preto, Brasil

Hendy Barbosa Santos

Faculdade de Artes do Paraná, Brasil

Humberto Costa

Universidade Federal do Paraná, Brasil

Igor Alexandre Barcelos Graciano Borges

Universidade de Brasília, Brasil

Inara Antunes Vieira Willerding

Universidade Federal de Santa Catarina, Brasil

Jaziel Vasconcelos Dorneles

Universidade de Coimbra, Portugal

Jean Carlos Gonçalves

Universidade Federal do Paraná, Brasil

Joao Adalberto Campato Junior

Universidade Brasil, Brasil

Jocimara Rodrigues de Sousa

Universidade de São Paulo, Brasil

Joelson Alves Onofre

Universidade Estadual de Santa Cruz, Brasil

Jónata Ferreira de Moura

Universidade São Francisco, Brasil

Jonathan Machado Domingues

Universidade Federal de São Paulo, Brasil

Jorge Eschriqui Vieira Pinto

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Jorge Luís de Oliveira Pinto Filho

Universidade Federal do Rio Grande do Norte, Brasil

Juliana de Oliveira Vicentini

Universidade de São Paulo, Brasil

Juliano Milton Kruger

Instituto Federal do Amazonas, Brasil

Juliano Pizzano Ayoub

Universidade Estadual de Ponta Grossa, Brasil

Julierme Sebastião Moraes Souza

Universidade Federal de Uberlândia, Brasil

Junior César Ferreira de Castro

Universidade de Brasília, Brasil

Katia Bruginski Mulik

Universidade de São Paulo, Brasil

Laionel Vieira da Silva

Universidade Federal da Paraíba, Brasil

Lauro Sérgio Machado Pereira

Instituto Federal do Norte de Minas Gerais, Brasil

Leonardo Freire Marino

Universidade do Estado do Rio de Janeiro, Brasil

Leonardo Pinheiro Mozdzenski

Universidade Federal de Pernambuco, Brasil

Letícia Cristina Alcântara Rodrigues

Faculdade de Artes do Paraná, Brasil

Lucila Romano Tragtenberg

Pontifícia Universidade Católica de São Paulo, Brasil

Lucimara Rett

Universidade Metodista de São Paulo, Brasil

Luiz Eduardo Neves dos Santos

Universidade Federal do Maranhão, Brasil

Maikel Pons Giralt

Universidade de Santa Cruz do Sul, Brasil

Manoel Augusto Polastrelí Barbosa

Universidade Federal do Espírito Santo, Brasil

Marcelo Nicomedes dos Reis Silva Filho

Universidade Estadual do Oeste do Paraná, Brasil

Márcia Alves da Silva

Universidade Federal de Pelotas, Brasil

Marcio Bernardino Sirino

Universidade Federal do Estado do Rio de Janeiro, Brasil

Marcos Pereira dos Santos

Universidad Internacional Iberoamericana del Mexico, México

Marcos Uzel Pereira da Silva

Universidade Federal da Bahia, Brasil

Marcus Fernando da Silva Praxedes

Universidade Federal do Recôncavo da Bahia, Brasil

Maria Aparecida da Silva Santandel

Universidade Federal de Mato Grosso do Sul, Brasil

Maria Cristina Giorgi

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brasil

Maria Edith Maroca de Avelar

Universidade Federal de Ouro Preto, Brasil

Marina Bezerra da Silva

Instituto Federal do Piauí, Brasil

Marines Rute de Oliveira

Universidade Estadual do Oeste do Paraná, Brasil

Maurício José de Souza Neto

Universidade Federal da Bahia, Brasil

Michele Marcelo Silva Bortolai

Universidade de São Paulo, Brasil

Mônica Tavares Orsini

Universidade Federal do Rio de Janeiro, Brasil

Nara Oliveira Salles

Unversidade do Estado do Rio de Janeiro, Brasil

Neide Araujo Castilho Teno

Universidade Estadual de Mato Grosso do Sul, Brasil

Neli Maria Mengalli

Pontifícia Universidade Católica de São Paulo, Brasil

Patricia Biegging

Universidade de São Paulo, Brasil

Patricia Flavia Mota

Universidade do Estado do Rio de Janeiro, Brasil

Patrícia Helena dos Santos Carneiro

Unversidade Federal de Rondônia, Brasil

Rainei Rodrigues Jadejiski

Universidade Federal do Espírito Santo, Brasil

Raul Inácio Busarello

Universidade Federal de Santa Catarina, Brasil

Raymundo Carlos Machado Ferreira Filho

Universidade Federal do Rio Grande do Sul, Brasil

Ricardo Luiz de Bittencourt

Universidade do Extremo Sul Catarinense, Brasil

Roberta Rodrigues Ponciano

Universidade Federal de Ubertândia, Brasil

Robson Teles Gomes

Universidade Católica de Pernambuco, Brasil

Rodiney Marcelo Braga dos Santos

Universidade Federal de Roraima, Brasil

Rodrigo Amancio de Assis

Universidade Federal de Mato Grosso, Brasil

Rodrigo Sarruge Molina

Universidade Federal do Espírito Santo, Brasil

Rogério Rauber

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Rosane de Fatima Antunes Obregon

Universidade Federal do Maranhão, Brasil

Samuel André Pompeo

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Sebastião Silva Soares

Universidade Federal do Tocantins, Brasil

Silmar José Spinardi Franchi

Universidade Federal de Santa Catarina, Brasil

Simone Alves de Carvalho

Universidade de São Paulo, Brasil

Simoni Urnau Bonfiglio

Universidade Federal da Paraíba, Brasil

Stela Maris Vaucher Farias

Universidade Federal do Rio Grande do Sul, Brasil

Tadeu João Ribeiro Baptista

Universidade Federal do Rio Grande do Norte, Brasil

Taiane Aparecida Ribeiro Nepomoceno

Universidade Estadual do Oeste do Paraná, Brasil

Taíza da Silva Gama

Universidade de São Paulo, Brasil

Tania Micheline Miorando

Universidade Federal de Santa Maria, Brasil

Tarcísio Vanzin

Universidade Federal de Santa Catarina, Brasil

Tascieli Feltrin

Universidade Federal de Santa Maria, Brasil

Tatiana da Costa Jansen

Serviço Nacional de Aprendizagem Comercial, Brasil

Tayson Ribeiro Teles

Universidade Federal do Acre, Brasil

Thiago Barbosa Soares

Universidade Federal do Tocantins, Brasil

Thiago Camargo Iwamoto

Universidade Estadual de Goiás, Brasil

Thiago Medeiros Barros

Universidade Federal do Rio Grande do Norte, Brasil

Tiago Mendes de Oliveira

Universidade Federal do Rio Grande do Norte, Brasil

Vanessa de Sales Marruche

Universidade Federal do Amazonas, Brasil

Vanessa Elisabete Raue Rodrigues

Universidade Estadual do Centro Oeste, Brasil

Vania Ribas Ulbricht
Universidade Federal de Santa Catarina, Brasil

Vinicius da Silva Freitas
Centro Universitário Vale do Cricaré, Brasil

Wellington Furtado Ramos
Universidade Federal de Mato Grosso do Sul, Brasil

Wellton da Silva de Fatima
Instituto Federal de Alagoas, Brasil

Wenis Vargas de Carvalho
Universidade Federal da Grande Dourados, Brasil

Yan Masetto Nicolai
Universidade Federal de São Carlos, Brasil

PARECERISTAS E REVISORES(AS) POR PARES

Avaliadores e avaliadoras Ad-Hoc

Alcidinei Dias Alves
Logos University International, Estados Unidos

Alessandra Figueiró Thornton
Universidade Luterana do Brasil, Brasil

Alexandre João Appio
Universidade do Vale do Rio dos Sinos, Brasil

Artur Pires de Camargos Júnior
Universidade do Vale do Sapucaí, Brasil

Bianka de Abreu Severo
Universidade Federal de Santa Maria, Brasil

Carlos Eduardo B. Alves
Universidade Federal do Agreste de Pernambuco, Brasil

Carlos Eduardo Damian Leite
Universidade de São Paulo, Brasil

Catarina Prestes de Carvalho
Instituto Federal Sul-Rio-Grandense, Brasil

Davi Fernandes Costa
Secretaria Municipal de Educação de São Paulo, Brasil

Denilson Marques dos Santos
Universidade do Estado do Pará, Brasil

Domingos Aparecido dos Reis
Must University, Estados Unidos

Edson Vieira da Silva de Camargos
Logos University International, Estados Unidos

Edwins de Moura Ramires
Serviço Nacional de Aprendizagem Comercial, Brasil

Elisiene Borges Leal
Universidade Federal do Piauí, Brasil

Elizabete de Paula Pacheco
Universidade Federal de Ubertândia, Brasil

Elton Simomukay
Universidade Estadual de Ponta Grossa, Brasil

Francisco Geová Goveia Silva Júnior
Universidade Potiguar, Brasil

Indiamaris Pereira
Universidade do Vale do Itajaí, Brasil

Jacqueline de Castro Rimá
Universidade Federal da Paraíba, Brasil

Jonas Lacchini
Pontifícia Universidade Católica de Minas Gerais, Brasil

Lucimar Romeu Fernandes
Instituto Politécnico de Bragança, Brasil

Marcos de Souza Machado
Universidade Federal da Bahia, Brasil

Michele de Oliveira Sampaio
Universidade Federal do Espírito Santo, Brasil

Nívea Consuelo Carvalho dos Santos
Serviço Nacional de Aprendizagem Comercial, Brasil

Pedro Augusto Paula do Carmo
Universidade Paulista, Brasil

Rayner do Nascimento Souza
Serviço Nacional de Aprendizagem Comercial, Brasil

Samara Castro da Silva
Universidade de Caxias do Sul, Brasil

Sidney Pereira Da Silva
Stockholm University, Suécia

Suêlen Rodrigues de Freitas Costa
Universidade Federal do Espírito Santo, Brasil

Thais Karina Souza do Nascimento
Instituto de Ciências das Artes, Brasil

Viviane Gil da Silva Oliveira
Universidade Federal do Amazonas, Brasil

Walmir Fernandes Pereira
Miami University of Science and Technology, Estados Unidos

Weyber Rodrigues de Souza
Pontifícia Universidade Católica de Goiás, Brasil

William Roslindo Paranhos
Universidade Federal de Santa Catarina, Brasil

Parecer e revisão por pares

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Pimenta Cultural, bem como revisados por pares, sendo indicados para a publicação.

*A informação se faz nas
práticas. O sentido se faz
no uso: texto e contexto.*

SUMÁRIO

Fabiano Couto Corrêa da Silva

Prefácio 14

Fabiano Couto Corrêa da Silva

Introdução 18

EIXO 1

**FUNDAÇÕES EPISTÊMICAS
E CRÍTICAS..... 29**

CAPÍTULO 1

Rodrigo Silva Caxias de Sousa

Gláucia Aparecida Vaz

Elisângela Gomes

Letícia Pereira de Souza

**Metodologias decoloniais para estudos
étnico-raciais na Ciência da Informação 31**

CAPÍTULO 2

Jackson da Silva Medeiros

**Problematizações e precauções
de método em ciência da informação.....69**

CAPÍTULO 3

Thiago Henrique Bragato Barros

**Análise de domínio enquanto um aporte
teórico a ciência da informação.....88**

EIXO 2

**ABORDAGENS MÉTRICAS
E AVALIATIVAS..... 113**

CAPÍTULO 4

Vanessa Inácio de Souza

Jussara Borges

Psicometria:

metodologia para avaliação em ciência da informação..... **115**

CAPÍTULO 5

Ana Maria Mielniczuk de Moura

Fernanda Bochi

Janaina Lais Pacheco Lara Morandin

Maria Cláudia Cabrini Grácio

Métodos de pesquisa aplicáveis

aos estudos de patentes na ciência

da informação..... 142

CAPÍTULO 6

Thais Dias Medeiros

Nicole Tirello Acquolini

Samile Andréa de Souza Vanz

Metodologias para estudar

rankings universitários..... 185

CAPÍTULO 7

Sônia Elisa Caregnato

Pedro Henrique da Silva Rodrigues

Carolina Carvalho Sena

Geise Ribeiro da Silva

Métodos de pesquisa

em comunicação científica:

estudos de impacto e visibilidade por meio de citações abertas..... **219**

CAPÍTULO 8

Rene Faustino Gabriel Junior

**Estudos métricos da informação
no Brasil:**

o uso da Brapci..... 248

CAPÍTULO 9

Valdir José Morigi

Luis Fernando Massoni

**A narratologia aplicada aos estudos
sobre informação, memória social**

e representações culturais..... 283

EIXO 3

PARADIGMAS DIGITAIS

E ABERTOS..... 309

CAPÍTULO 10

Fabiano Couto Corrêa da Silva

Metodologias para a ciência aberta:

métricas, governança e práticas na ciência da informação 311

CAPÍTULO 11

Rafael Port da Rocha

Digitalização e preservação digital:

abordagens metodológicas 353

CAPÍTULO 12

Caterina Marta Groposo Pavão

Ana Julia Lopes

Lucia Helena Cunha Vidal

Dados que falam por si:

avaliação da qualidade dos dados para uma pesquisa
transparente e reprodutível 379

CAPÍTULO 13

Moisés Rockembach

Priscilla Van Even

Jonas Ferrigolo Melo

Luciana Monteiro-Krebs

Metodologias de co-design em ciência

da informação e comunicação 431

Sobre os autores e as autoras 471

Índice remissivo 484

PREFÁCIO

"Vivemos em uma sociedade profundamente dependente da ciência e da tecnologia, em que quase ninguém sabe algo sobre ciência e tecnologia." Trinta e cinco anos depois, em um ecossistema de circulação acelerada e fragmentada, fértil para a desinformação, a advertência de Carl Sagan (1990)¹ ganha urgência: comunicar ciência com precisão e contexto é antídoto público e condição para decisões coletivas responsáveis. É também por isso que a agenda da ciência aberta se afirma como resposta concreta: abrir dados, métodos e códigos não é apenas eficiência acadêmica, mas estratégia de inteligibilidade e verificabilidade do conhecimento de que dependemos todos os dias.

No campo da Ciência da Informação, esse movimento é particularmente marcante. Navegamos entre desafios e oportunidades próprios da era digital. É nesse contexto de conexão permanente e transformação acelerada que se insere esta obra, apresentando metodologias de pesquisa em diversas vertentes e abrindo caminhos para compreender o papel da informação na construção do conhecimento científico e na vida social contemporânea. A rápida digitalização do conhecimento e as demandas por transparência, acessibilidade e integridade pressionam pesquisadores e profissionais da informação a revisitar práticas e métodos. Este livro responde a tais demandas ao oferecer um panorama detalhado de abordagens metodológicas atualizadas e testadas na prática.

1 O artigo "Why We Need to Understand Science", de autoria de Carl Sagan, foi publicado na revista *Skeptical Inquirer*, v. 14, n. 3 (primavera de 1990), p. 264: "We live in a society exquisitely dependent on science and technology, in which hardly anyone knows anything about science and technology.". Disponível em: www.carlsagan.com

SUMÁRIO

Mais do que um compêndio de técnicas, a obra fomenta uma cultura investigativa comprometida com a complexidade dos fenômenos informacionais, com a pluralidade epistêmica que os caracteriza e com a responsabilidade social do fazer científico. Não se limita ao “como fazer”; convida ao pensar crítico que sustenta cada decisão metodológica e eleva o debate no campo.

Por essência, a Ciência da Informação é interdisciplinar: integra aportes das ciências sociais, da computação, da comunicação, da administração e de outras áreas. Dessa confluência emerge uma pluralidade de métodos que responde à variedade de problemas que investigamos, da produção e da organização à mediação, ao uso e à preservação da informação. A trajetória do campo mostra um movimento de sofisticação contínua, de abordagens inicialmente descritivas e voltadas a desafios práticos para um repertório amplo e articulado, em que convivem paradigmas positivistas, pós-positivistas, interpretativos, críticos e pragmáticos. Essa pluralidade não é fraqueza; é força estratégica para encarar objetos multifacetados com a profundidade que nenhum método isolado alcançaria.

Nesse horizonte, a adequação metodológica é princípio orientador: cada pergunta de pesquisa demanda um desenho compatível com seus objetivos e com o referencial teórico que a ampara. A obra recusa universalismos e monismos e advoga por um pragmatismo orientado pelo problema, com rigor entendido como coerência entre problema, teoria e método, e não como rigidez. Com a digitalização intensa e a profusão de dados, essa coerência se articula à transparência e à reprodutibilidade. A ciência aberta desloca o foco do “resultado” para o processo, exige abertura de dados, métodos e códigos, e favorece cooperação, escrutínio e verificação independente.

Essa visão se materializa na estrutura do livro ao longo de quatro eixos que dialogam entre si. O primeiro, dedicado às fundações epistêmicas e críticas, tensiona bases filosóficas e éticas do fazer metodológico em CI. Ao acolher perspectivas decoloniais,

SUMÁRIO

desnaturaliza a ideia de “epistemologia única” e mostra que escolhas de método são também escolhas políticas, pois implicam reconhecer vozes, histórias e contextos. Abordagens como Escrivência, Imagens de Controle e Análise Episódica operam como ferramentas para desvelar e reconfigurar dinâmicas de poder e formas de silenciamento na produção do conhecimento. Em paralelo, a reflexão sobre o próprio objeto “informação” recusa a neutralidade: informação é fenômeno situado e discursivo, cujo sentido se faz no uso e nas formas de vida que o sustentam. Daí decorre um ponto central: metodologia não é apêndice, integra a constituição do objeto investigado.

O segundo eixo desloca a discussão para os modos como estruturamos e atribuímos sentido à informação. Ao trabalhar com narratologia, um método das humanidades incorporado criticamente pela CI, explora-se a análise de informação, memória social e representações culturais em suportes textuais, orais, visuais e digitais. Em tempos de fragmentação e micronarrativas, a narratologia ajuda a apreender lógicas, temporalidades e ideologias subjacentes às mensagens, mostrando que organizar informação é também interpretar, e que descrição, indexação e representação são decisões com efeitos cognitivos e sociais. Além disso, o volume dedica capítulos aos Sistemas de Organização do Conhecimento, tanto em seu contexto quanto em aplicações (*Metodologias de Pesquisa no Contexto dos Sistemas de Organização do Conhecimento* e *Metodologias de Pesquisa na Organização do Conhecimento*), e a uma epistemologia aplicada à área (*Epistemologia da Ciência da Informação: Métodos e Perspectivas de Pesquisa*), reforçando que organizar depende de fundamentos teóricos explícitos e de escolhas metodológicas coerentes.

O terceiro eixo evidencia o potencial analítico da CI ao articular instrumentos quantitativos e computacionais com leitura crítica. A psicometria oferece meios para mensurar construtos latentes em competências infocomunicacionais, com ênfase em validade e fidedignidade. A análise de patentes amplia o repertório com patentometria, prospecção, citações e redes, convertendo um corpus técnico em

SUMÁRIO

inteligência informacional. As metodologias para estudar rankings universitários desnudam convenções e interesses que moldam métricas tidas como objetivas. Somam-se os métodos de pesquisa em comunicação científica por meio de citações abertas, que fortalecem transparência, reprodutibilidade e novas leituras de impacto. A síntese aplicada dos estudos métricos no Brasil, com Brapci, Lattes e outras bases, reforça que contar não basta: é preciso qualificar metadados, contextualizar séries históricas e interpretar resultados.

O quarto eixo responde aos desafios da era dos dados e da ciência aberta. Reprodutibilidade, transparência e colaboração tornam-se exigências de método. A preservação digital aparece como política e como técnica, condição para a perenidade do registro científico e para a memória institucional. A qualidade dos dados é tratada como fundamento da inferência confiável. E o co-design aproxima pesquisa e sociedade, envolvendo múltiplas partes interessadas na construção de problemas, hipóteses, instrumentos e soluções, fortalecendo a relevância pública da investigação e uma ética do cuidado com contextos e comunidades.

Em conjunto, os capítulos compõem um livro que é, ao mesmo tempo, mapa e convite. Mapa, porque oferece rotas para quem inicia e atalhos críticos para quem já percorre o campo; convite, porque chama a revisitar pressupostos e escolhas, a experimentar métodos, a documentar percursos e a partilhar processos. A mensagem é clara: a força da Ciência da Informação está menos em adotar um método e mais em escolhê-lo bem, à luz do problema, do quadro teórico, do contexto e do compromisso social da pesquisa. Que esta cartografia ajude pesquisadoras e pesquisadores, em qualquer estágio, a navegar um terreno móvel com coerência, imaginação e responsabilidade, consolidando a CI como campo plural, inovador e socialmente relevante.

Fabiano Couto Corrêa da Silva

INTRODUÇÃO

Fabiano Couto Corrêa da Silva

A Ciência da Informação (CI) caracteriza-se por uma ampla diversidade de abordagens metodológicas. Por sua natureza interdisciplinar, o campo integra perspectivas das ciências sociais, computação, comunicação, administração e outras áreas, o que se reflete em uma pluralidade de métodos de pesquisa. Essa diversidade metodológica é resultado da complexidade dos fenômenos informacionais estudados e das múltiplas vertentes teóricas e epistemológicas que convivem na área, ao longo da trajetória da área. Desse modo, compreender a importância de adotar metodologias adequadas em CI requer analisar como diferentes paradigmas de pesquisa demandam abordagens específicas, bem como reconhecer a relação intrínseca entre ciência, comunicação científica, dados e práticas metodológicas no desenvolvimento do campo. Nesta coletânea, os capítulos foram escritos por docentes do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS), em colaboração com estudantes e pesquisadores(as) parceiros(as), ou de forma individual refletindo um processo formativo e colaborativo que também é objeto desta discussão.

A escolha da metodologia de pesquisa é um elemento essencial para a confiabilidade e a validade de qualquer investigação científica. No campo da CI, essa escolha reveste-se de particular importância dado o caráter multifacetado do objeto de estudo: a informação em seus contextos de produção, organização, mediação, uso e preservação. Segundo Ullah e Ameen (2018), a utilização de metodologias adequadas é componente necessário para garantir a robustez de uma pesquisa, influenciando diretamente a qualidade dos resultados obtidos e sua aceitação pela comunidade científica.

SUMÁRIO

Em uma revisão sistemática de métodos em Biblioteconomia e CI, esses autores identificaram que pesquisas empíricas de cunho descritivo e quantitativo ainda dominam uma parte considerável da literatura internacional, com destaque histórico para métodos de levantamento (surveys) e uso frequente de estatísticas descritivas na análise de dados (Ullah; Ameen, 2018). Esse panorama indica uma tradição forte de pesquisas aplicadas e quantitativas no campo, especialmente em países centrais, o que se relaciona a uma fase de desenvolvimento da CI voltada a medições objetivas e mensuração de fenômenos informacionais.

No entanto, a realidade metodológica da CI é mais complexa e vem se ampliando. Estudos de alcance internacional sobre tendências metodológicas mostram que, ao longo das décadas, houve evolução das estratégias de pesquisa empregadas na área, incluindo um gradual aumento de sofisticação metodológica e diversificação de métodos (Risso, 2016; Ullah; Ameen, 2018). Historicamente, conforme aponta Risso (2016, p. 74), o desenvolvimento inicial da CI esteve estreitamente ligado a metodologias de natureza descritiva, voltadas a atender desafios da prática profissional (por exemplo, estudos sobre catalogação, uso de bibliotecas, desenvolvimento de coleções, recuperação da informação) por meio de estratégias empíricas. Com o tempo, contudo, pesquisadores da informação passaram a adotar métodos mais variados e rigorosos, incorporando técnicas provenientes de outras disciplinas e explorando novos desenhos de pesquisa. Essa expansão metodológica acompanha as transformações do próprio campo: à medida que a CI consolidou seu estatuto científico e ampliou seu escopo temático, emergiram diferentes paradigmas de pesquisa, como o positivista, o pós-positivista, o interpretativo, o crítico e o pragmático, entre outros, cada qual orientando preferências metodológicas distintas.

É fundamental reconhecer que por trás de cada escolha metodológica há pressupostos epistemológicos. Um estudo realizado no contexto brasileiro por Borges e Maricato (2022) evidenciou

SUMÁRIO

a predominância de abordagens qualitativas em grande parte das investigações recentes daquele contexto, muitas delas ancoradas em revisões bibliográficas e pesquisas descritivas de natureza aplicada. Ao examinarem 134 dissertações e teses de um programa de pós-graduação em CI, os autores encontraram que a maioria indicava métodos qualitativos e de revisão de literatura como centrais, seguidos por pesquisas descritivas e aplicadas (Borges; Maricato, 2022). Esse dado sugere que, ao menos em determinados contextos acadêmicos, na realidade brasileira, há forte valorização de métodos interpretativos e teóricos, possivelmente associados a vertentes epistemológicas construtivistas, críticas ou humanísticas da CI. Contrasta, em parte, com levantamentos internacionais que apontavam peso considerável de métodos quantitativos; a diferença pode refletir tradições acadêmicas locais e objetos de pesquisa específicos. De todo modo, o quadro geral aponta para uma convivência de múltiplos enfoques metodológicos na Ciência da Informação – convivência esta que é saudável e necessária, dado o espectro variado de problemas investigados.

Nessa perspectiva, a adequação metodológica torna-se um ponto nevrálgico: cada problema de pesquisa em CI demanda um desenho metodológico compatível com suas perguntas, objetivos e referencial teórico. Uma investigação sobre comportamento informacional de usuários, por exemplo, poderá adotar entrevistas em profundidade, observação etnográfica ou análise fenomenológica para captar a subjetividade e o contexto das práticas informacionais, abordagens coerentes com paradigmas interpretativos. Por outro lado, um estudo voltado a mensurar a produção científica de uma área provavelmente recorrerá a métodos quantitativos, como análises bibliométricas ou cientométricas, inseridas em um arcabouço pós-positivista ou funcionalista que privilegia a objetividade e a quantificação. Adotar uma metodologia condizente com a pergunta de pesquisa e ao marco teórico, garantindo consistência interna ao estudo. Quando há alinhamento, os resultados tendem a ser mais

SUMÁRIO

significativos e comunicáveis; do contrário, corre-se o risco de chegar a conclusões pouco válidas ou incoerentes com o que se pretendia investigar.

A relação entre ciência, comunicação científica, dados e práticas metodológicas merece destaque especial nesse contexto, pois a ciência contemporânea enfatiza cada vez mais a transparência metodológica e a reprodutibilidade dos estudos, fatores diretamente vinculados à forma como os resultados podem ser comunicados e validados pela comunidade. Um delineamento metodológico claro facilita a descrição precisa dos procedimentos e permite que outros pesquisadores compreendam e eventualmente repitam o estudo, ao passo que fundamentos frágeis comprometem a confiança e dificultam a circulação do conhecimento. Esse cenário é intensificado pela era dos dados intensivos: big data, bases bibliográficas, registros digitais, altmétricos e outras fontes remodelam tanto os objetos quanto às práticas da Ciência da Informação, demandando competências híbridas que integram estatística, ciência de dados, algoritmos e fundamentos teóricos de análise de redes e representação do conhecimento. Investigações sobre coautoria, ontologias ou visualizações complexas ilustram como metodologias tradicionais e emergentes passaram a coexistir no cotidiano da pesquisa em CI. A comunicação científica aberta acrescenta uma dimensão decisiva a esse quadro: a publicação de protocolos, o compartilhamento de dados brutos, os repositórios de preprints e o open peer review tornam-se práticas correntes, ampliando colaboração e escrutínio independente, mas também exigindo padronização, qualidade e reprodutibilidade. Como observam Albagli, Clinio e Raychtock (2014), a ciência aberta envolve múltiplas correntes interpretativas e tipos de ação, do acesso aberto à literatura a práticas cidadãs de produção de conhecimento, interpelando pesquisadores da CI a redesenhar suas metodologias de modo mais transparente, colaborativo e alinhado aos princípios de abertura.

SUMÁRIO

Nesse panorama, podemos observar que cada subárea da Ciência da Informação demanda abordagens metodológicas específicas e enfrenta questões próprias quanto à escolha dos métodos. Vejamos algumas dessas subáreas e como elas exemplificam a diversidade de métodos no campo:

Estudos métricos da informação ocupam lugar tradicional na CI, ao quantificar e analisar a informação registrada por meio da bibliometria, cientometria, informetria e altmetria. Embora ancorados em métodos quantitativos e modelos estatísticos, esses estudos vão além de números: revelam como a ciência se organiza, circula e é avaliada. O desafio contemporâneo está em evitar a ilusão de neutralidade dos indicadores e compreender que métricas são também construções sociais, marcadas por escolhas e exclusões. A reflexão metodológica, nesse caso, implica não apenas aperfeiçoar técnicas de mensuração, mas problematizar o que os números silenciam e de que maneira podem ser combinados a análises qualitativas para captar dimensões mais amplas da produção científica.

Organização do Conhecimento (OC), por sua vez, evidencia outra face da CI: o esforço de classificar, representar e estruturar conceitos que orientam a mediação da informação. Pesquisas nesse campo, predominantemente qualitativas e teóricas, lidam com sistemas abstratos, linguagens documentárias e modelos conceituais/computacionais. No entanto, os métodos aqui empregados não são apenas instrumentos técnicos, mas escolhas epistemológicas que definem como os sujeitos, as culturas e os saberes são representados — e, portanto, incluídos ou silenciados. A reflexão metodológica nesse domínio convida a explicitar tais escolhas, reconhecendo que organizar conhecimento é também intervir nos modos de pensar e narrar o mundo.

Mediação da informação e uso da informação trazem ao centro a dimensão social, cultural e política do acesso ao conhecimento. As metodologias empregadas frequentemente qualitativas,

SUMÁRIO

participativas e críticas permitem captar a complexidade das práticas informacionais em bibliotecas, arquivos ou espaços digitais. Mais do que mapear comportamentos, a pesquisa neste campo revela relações de poder, desigualdades de acesso e disputas de sentido. A escolha metodológica, aqui, reflete uma tomada de posição: ao adotar narrativas, memórias ou metodologias decoloniais, pesquisadores reconhecem que a mediação não é neutra, mas um ato ético e político de construção de significados.

Curadoria digital e preservação da informação traduzem o encontro entre técnica e cuidado. O foco em procedimentos de manutenção e perenização de acervos digitais demanda métodos rigorosos, técnicos e interdisciplinares, mas a reflexão metodológica vai além da dimensão operacional. Preservar dados e registros é também preservar memórias, histórias e identidades, o que coloca o pesquisador diante de escolhas éticas e institucionais. Nessa perspectiva, metodologias aplicadas à curadoria digital não apenas asseguram confiabilidade técnica, mas se constituem como práticas sociais que moldam a forma como comunidades futuras terão acesso ao conhecimento do presente.

Ciência aberta e ciência de dados aplicadas à CI representam, talvez, a expressão mais contemporânea desse debate. Mais do que um conjunto de práticas técnicas como disponibilização de dados, publicação de preprints ou revisão aberta, a ciência aberta questiona os próprios fundamentos da produção científica: quem participa, quem acessa e quem valida o conhecimento. Nesse campo, as metodologias não apenas descrevem práticas, mas precisam incorporar os princípios da abertura, tornando-se elas mesmas mais transparentes, colaborativas e reproduzíveis. O desafio está em equilibrar exigências de padronização e rigor com a necessidade de respeitar contextos, diversidades e limites éticos.

Tomadas em conjunto, essas subáreas ilustram que a metodologia em Ciência da Informação não se reduz a técnicas ou

SUMÁRIO

protocolos: ela é sempre uma escolha situada, que envolve compromissos teóricos, éticos e políticos. Ao revisitar as práticas metodológicas sob diferentes ângulos, da quantificação à mediação, da representação à abertura, este livro convida a reconhecer que investigar informação é, ao mesmo tempo, investigar as condições de produção de conhecimento em nossa sociedade. Assim, a diversidade que percorremos até aqui revela algo essencial: métodos não são ferramentas neutras; são compromissos teóricos, éticos e políticos que orientam o olhar do pesquisador.

Nenhuma escolha é neutra. Ela carrega pressupostos sobre o que é conhecimento válido, qual a melhor forma de obter evidências e até que papel o pesquisador assume diante do fenômeno. Por isso, pesquisadores dedicam atenção à fundamentação metodológica de seus trabalhos, justificando por que um estudo de caso qualitativo, por exemplo, é apropriado para explorar uma iniciativa de biblioteca escolar, ou por que uma análise estatística de séries temporais faz sentido para examinar a evolução da produção científica numa área. Essa justificativa não é mero formalismo: ela demonstra alinhamento entre as perguntas formuladas, o referencial teórico adotado e os meios escolhidos para obter respostas. Em suma, demonstra coerência científica.

Nos últimos anos, também se observa um movimento de integração metodológica na Ciência da Informação, com maior utilização de métodos mistos (mixed methods). Embora ainda corresponda a parcela relativamente pequena da produção (menos de 0,5% segundo revisão de Hayman e Smith (2020) para 2008–2018), há um aumento gradual de estudos que combinam técnicas quantitativas e qualitativas num mesmo projeto. A adoção de métodos mistos pode ser extremamente frutífera em CI, pois muitos fenômenos informacionais têm tanto dimensões mensuráveis quanto interpretativas. Por exemplo, uma pesquisa sobre impacto de repositórios institucionais pode levantar dados quantitativos de acesso e downloads (método quantitativo) e também conduzir entrevistas com pesquisadores

SUMÁRIO

depositantes para compreender suas motivações e obstáculos (método qualitativo). Ao triangular métodos, obtém-se uma visão mais holística do objeto de estudo, ao mesmo tempo mitigando limitações inerentes a cada abordagem isoladamente. No entanto, o uso de métodos mistos exige cuidado metodológico dobrado: é preciso garantir que ambas as partes, quantitativa e qualitativa, sejam conduzidas com rigor e que haja uma integração coerente dos resultados na interpretação final. Esse tipo de reflexão vem ganhando espaço, indicando maior maturidade metodológica do campo.

Diante de tudo que foi exposto, fica claro que a diversidade de abordagens metodológicas na Ciência da Informação é tanto uma necessidade quanto um trunfo. Nenhum único método conseguiria abarcar a complexidade dos fenômenos que estudamos – da circulação global do conhecimento científico às interações locais de um usuário com um sistema de busca, das estruturas conceituais que organizam bibliotecas às dinâmicas de poder informacional em contextos decoloniais. Assim, a abertura do campo a diferentes vertentes teórico-metodológicas é sinal de vitalidade e capacidade de adaptação. Contudo, essa pluralidade demanda do pesquisador uma postura crítica e informada na seleção e aplicação de metodologias. Adotar metodologias adequadas, como requerido, implica atualizar-se sobre novos métodos que surgem, conhecer as limitações de cada técnica, combinar abordagens quando pertinente e, sobretudo, assegurar que o caminho metodológico escolhido responde às perguntas propostas de modo ético e eficaz.

É nesse sentido que a reflexão metodológica adquire um papel central: mais do que um conjunto de técnicas e procedimentos operacionais, a metodologia de pesquisa em Ciência da Informação deve ser vista como um componente estratégico da investigação, que orienta o pensar e o fazer científico. Cada vertente teórica, seja positivista, interpretativa, crítica, pós-moderna, traz consigo não apenas uma visão de mundo, mas também preferências metodológicas (a experimentação, a estatística, a hermenêutica, a genealogia, etc.).

SUMÁRIO

Reconhecer essa vinculação evita simplificações e ecletismos ingênuos, nos quais métodos são aplicados de forma deslocada de seus contextos paradigmáticos. Ao mesmo tempo, o diálogo interdisciplinar que caracteriza a CI permite apropriações criativas de métodos de outros campos, adaptando-os aos objetivos informacionais. Foi assim que a análise de redes (originada na sociologia e matemática) passou a ser aplicada na bibliometria para redes de citações, ou que a análise de conteúdo de Bardin (2006), típica das ciências sociais, se tornou comum em estudos de usuários e mediação na CI. Tais trocas metodológicas enriquecem o campo, desde que realizadas com o devido rigor.

Um olhar atento às literaturas recentes reforça essas constatações. Revisões e mapeamentos publicados nos últimos anos têm oferecido panoramas do uso de métodos na CI, indicando tanto tendências quanto lacunas. Por exemplo, Barros e Vital (2019) identificaram lacunas na explicitação metodológica nos trabalhos de Organização do Conhecimento no Brasil, sugerindo a necessidade de maior clareza e formação metodológica para pesquisadores dessa subárea. Borges e Maricato (2022) enfatizaram a importância de sistematizar as características metodológicas das pesquisas dos programas de pós-graduação, de modo a aprimorar a formação dos novos pesquisadores e fortalecer a identidade científica do campo. No âmbito internacional, Ullah e Ameen (2018) apontaram a falta de taxonomias de métodos em boa parte dos artigos analisados, o que dificulta a comparação e cumulação de conhecimento sobre práticas metodológicas. Essas análises críticas servem de alerta: apesar dos avanços, ainda há muito espaço para aprimorar a consciência metodológica dos pesquisadores em Ciência da Informação.

Em resposta a esse cenário, obras coletivas, como o livro em que este texto se insere, cumprem um papel fundamental ao oferecer uma cartografia crítica dos caminhos metodológicos possíveis na área. Mais do que um manual de técnicas, trata-se de fomentar uma cultura investigativa comprometida com a complexidade dos fenômenos

SUMÁRIO

informacionais, com a pluralidade epistêmica e com a responsabilidade social da ciência. Ao reunir experiências diversificadas, ancoradas em contextos reais de pesquisa, obras assim demonstram na prática como diferentes metodologias podem ser aplicadas e adaptadas para iluminar facetas distintas da informação. Espera-se, com isso, inspirar tanto pesquisadores iniciantes quanto experientes a reavaliar seus percursos metodológicos, experimentar novas possibilidades investigativas e refletir criticamente sobre as bases de suas escolhas.

Os desafios da diversidade metodológica na Ciência da Informação revelam não apenas a variedade de técnicas disponíveis, mas sobretudo a necessidade de escolhas conscientes e situadas. Adotar metodologias adequadas significa reconhecer a pluralidade de vertentes teóricas, epistemológicas e aplicadas que atravessam o campo, compreendendo que cada abordagem carrega consigo compromissos e implicações. Inserida em uma sociedade informacional em constante transformação, a CI evolui juntamente com suas metodologias: incorpora inovações tecnológicas, responde às demandas de transparência e rigor e, ao mesmo tempo, dialoga com problemas sociais emergentes que tensionam suas práticas.

REFERÊNCIAS

ALBAGLI, S.; CLINIO, A.; RAYCHTOCK, S. Ciência aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, p. 434–450, 2014.

ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. (Org.). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015.

ARAÚJO, C. A. A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11–32, 2006.

BARROS, C. M. de; VITAL, L. P. Abordagens metodológicas das pesquisas em organização e representação do conhecimento no contexto brasileiro. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 17, 2019.

SUMÁRIO

BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 2006.

BORGES, V. J.; MARICATO, J. M. Reflexões sobre as características metodológicas da pesquisa científica em Ciência da Informação. **Informação & Informação**, Londrina, v. 27, n. 3, p. 473–496, 2022.

HAYMAN, R.; SMITH, E. E. **Mixed Methods Research in Library and Information Science: A Methodological Review**. Evidence Based Library and Information Practice, v. 15, n. 1, p. 106–125, 2020. DOI: 10.18438/ebliip29648.

HIGGINS, S. Digital curation: the emergence of a new discipline. **International Journal of Digital Curation**, v. 6, n. 2, p. 78–88, 2011.

HJØRLAND, B. Domain analysis in information science: eleven approaches — traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422–462, 2002. DOI: 10.1108/00220410210431136.

KILOMBA, G. **Memórias da Plantação**: episódios de racismo cotidiano. 2. ed. Rio de Janeiro: Cobogó, 2019.

MUELLER, S. P. M. **Métodos para a pesquisa em Ciência da Informação**. Brasília: Thesaurus, 2007.

RISSO, V. G. Research methods used in library and information science during the 1970–2010. **New Library World**, v. 117, n. 1/2, p. 74–93, 2016.

ULLAH, A.; AMEEN, K. Account of methodologies and methods applied in LIS research: a systematic review. **Library & Information Science Research**, v. 40, n. 1, p. 53–60, 2018.

Eixo

1

FUNDAÇÕES EPISTÊMICAS E CRÍTICAS

A investigação dos alicerces filosóficos e críticos da Ciência da Informação define este eixo, no qual se questionam os métodos de pesquisa e se propõem novas abordagens, como os estudos decoloniais ou análise de domínio para uma compreensão mais profunda e socialmente consciente do campo.

1

Rodrigo Silva Caxias de Sousa
Gláucia Aparecida Vaz
Elisângela Gomes
Letícia Pereira de Souza

METODOLOGIAS DECOLONIAIS PARA ESTUDOS ÉTNICO-RACIAIS NA CIÊNCIA DA INFORMAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.1

SUMÁRIO

COMO CITAR:

SOUSA, Rodrigo Silva Caxias de; VAZ, Gláucia Aparecida; GOMES, Elisângela; SOUZA, Letícia Pereira de. Metodologias de pesquisa em estudos de decolonialidade na Ciência da Informação. In: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

A instituição e adoção de epistemologias decoloniais no âmbito da Ciência da Informação no Brasil vem se constituindo como uma perspectiva assumida recentemente por estudiosos da área, em razão de reflexões desenvolvidas a partir de disputas entre coletividades científicas; articuladas aos tensionamentos e influências do campo de estudos qualitativos; imbricadas a movimentos sociais e políticos. No que tange às metodologias decoloniais, pesquisadoras (es) de espaços de produção periférica do conhecimento têm estabelecido compreensões que considerem a experiência de vida individual e em comunidade; manifestada através da oralidade, nas práticas e nos modos de fazer; entendendo que a produção de conhecimento se efetiva justamente através do compartilhamento de saberes; sem conceber o conhecimento científico como um saber maior, acima de outros saberes. Ainda que historicamente a noção do atributo de conhecimento válido e autorizado tenha sido atribuída à ciência, compreensões dessa natureza necessitam ser ultrapassadas, relativizando a lógica de uma “epistemologia única”. A escritora, teórica, psicóloga e artista interdisciplinar Grada Kilomba atenta sobre o significado de epistemologia.

SUMÁRIO

Por favor, deixem-me lembrar-lhes o que significa o termo epistemologia. O termo é composto pela palavra grega *episteme*, que significa conhecimento, e *logos*, que significa ciência. Epistemologia é, então, a ciência da aquisição de conhecimento, que determina: (a). (os temas) quais temas ou tópicos merecem atenção e quais questões são dignas de serem feitas com o intuito de produzir conhecimento verdadeiro; (b). (os paradigmas) quais narrativas e interpretações podem ser usadas para explicar um fenômeno, isto é, a partir de qual perspectiva o conhecimento verdadeiro pode ser produzido; (c). (os métodos) e quais maneiras e formatos podem ser usados para a produção de conhecimento confiável e verdadeiro. Epistemologia, como eu já havia dito, define não somente como, mas também quem produz conhecimento verdadeiro e em quem acreditarmos (Kilomba, 2019, p. 10-11).

Considerando que tanto os indivíduos aos quais são atribuídas relações de validação do conhecimento quanto seus métodos, paradigmas e problemas são fruto de embates políticos e de relações de poder movimentos epistemológicos periféricos, passaram, no âmbito dos estudos decoloniais, a considerar a inserção de saberes plurais como alternativa de compreensão de fenômenos em diferentes áreas do conhecimento. Essas concepções e possibilidades projetam a necessidade de posturas calcadas em movimentos epistemológicos transgressores, em busca da instituição de um “outro lugar epistemológico” (FANON, 2008); de acordo com constantes movimentos de “desobediência epistêmica” (MIGNOLO, 2008), que pretendem desvelar e desconstruir aspectos fundadores destacados por Kilomba (2016) da dinâmica de produção do conhecimento científico.

A noção de um outro lugar epistemológico (FANON, 2008) se constitui como referência que considera uma inversão do sentido centro periferia. Essa inversão, ao permitir que outros atores sejam vistos como protagonistas no processo de produção do conhecimento científico, se desdobra em uma prática científica articulada à

SUMÁRIO

realidade dos espaços periféricos. Inverte-se a relação do norte-sul global, mas sobretudo a compreensão de que as investigações científicas considerem objetos de estudo, temas, formas de concepção, composição e validação de conhecimento; a partir de atores que se encontrem em consonância com as necessidades e demandas das coletividades a partir das quais se constituíram subjetividades e identidades políticas. Pensada, contrariamente, a partir do viés da decolonialidade a perspectiva prescritiva dando lugar a uma concepção de ciência e produção do conhecimento que se constitui à luz de que as investigações possam encontrar-se definidas segundo a construção e interesse social coletivo, como forma de alcançar objetivos comuns; intencionando mudanças concretas articuladas à vida dos indivíduos aos quais se relacionam tais fenômenos.

Quanto à cinesia da desobediência epistêmica (MIGNOLO, 2008), à luz das epistemologias decoloniais e suas potenciais aplicações nos estudos étnicos-raciais na Ciência da Informação, é indispensável assumir um rechaço à objetividade e universalidade do conhecimento científico produzido, aceito até então de forma naturalizada. É necessário assumir que a ciência não deva ser compreendida como uma atividade neutra, mas sobretudo, como atividade social e política imbricada aos multifacetados processos de colonialidade que se articulam historicamente. Consideradas essas rearticulações cabe destacar a forma como a colonialidade se caracteriza. Na perspectiva adotada por Quijano (1992) a colonialidade é a perpetuação das relações de poder colonial advindas da estrutura do passado, alicerçada na racialização da produção e seus desdobramentos nas relações humanas e de capital-trabalho; no imaginário eurocêntrico como dispositivos concretos e simbólicos de coordenação e controle dos corpos e subjetividades racializados; e na hegemonia do “Estado-nação” que confere a noção de espaços periféricos. Nesses espaços são refinadas as relações de poder negligenciados saberes e violentados e violentadas subjetividades; sendo que essas relações são reforçadas mutuamente.

SUMÁRIO

Da perpetuação e rearticulação constante das relações coloniais vem surgindo outras perspectivas de métodos, compreendidos como encadeamento de saberes plurais e que projetam conexões situadas nos contextos de produção subalterna, ao considerarem atravessamentos advindos das violências e opressões coloniais; requerendo um olhar diligente acerca da imbricação entre movimentos epistemológicos, teoria, conceitos adotados e metodologias.

Merece destaque que neste capítulo esse nosso olhar se propõe a apresentar, a partir de teóricas decoloniais, três perspectivas metodológicas que possam ser utilizadas como referência para os estudos étnico-raciais na CI, quais sejam: a Escrivivência de Conceição Evaristo (Evaristo, 1996), as Imagens de Controle de Patrícia Hill Collins (Collins, 2009) e a Análise Episódica de Grada Kilomba (Kilomba, 2019).

A ESCRIVIVÊNCIA COMO CAMINHO EMANCIPATÓRIO

A escrita orientada pelas experiências de vida se dá pela afirmação do sujeito negro centrado na sua vivência, que a partir desse posicionamento, constrói e elabora a sua própria imagem cultural e formas próprias de narrar e analisar acontecimentos históricos e culturais. Esse entendimento tornou-se um conceito crítico-literário nomeado pela escritora Conceição Evaristo como escrevivência, que em sua dissertação de mestrado defendida em 1996, afirma: “A escrevivência do corpo negro é realizada não só pela apresentação física desse corpo em si, de seus movimentos religiosos ou lúdicos, mas também pela construção material que esse corpo opera, que esse corpo produz” (EVARISTO, 1996, p. 105).

SUMÁRIO

Apoiada na sua condição de mulher negra em um contexto de vida periférico, ao realizar sua produção literária, a autora propõe afirmação do “eu” enquanto sujeito social, portanto, suas referências carregam contextos culturais, sociais e linguísticos que traduzem sua experiência de vida coletivizada. Dessa forma, a escritora aponta para a invisibilidade da literatura produzida por escritoras (es) negras, além de questionar o cânone literário, constituído essencialmente por indivíduos brancos que historicamente elaboraram estereótipos reducionistas e negativos sobre a população negra.

Essa crítica se concentra principalmente no período Romântico da literatura brasileira, que no século XIX, contribuiu para a construção de uma identidade cultural do país, a partir de uma suposta valorização nacional por meio da exaltação da natureza, da presença africana e indígena, das narrativas históricas e de cunho regional. Por se tratar de um movimento literário relacionado à ascensão da burguesia, os escritores desse período foram intelectuais predominantemente homens brancos pertencentes a elite cultural. Condicionados à sua realidade social, em suas escritas, construíram um imaginário estereotipado sobre a população negra, por meio da construção de personagens sem história de vida, objetificados e reduzidos a sua condição racial e social.

Ao investigar a trajetória negra na literatura brasileira, Proença Filho (2024, p. 161) afirma que esta “[...] não escapa ao tratamento marginalizado que, desde as instâncias fundadoras, marca a etnia no processo de construção da nossa sociedade”. A figura da “mãe-preta”, por exemplo, aparece nas narrativas históricas e literárias sobre o período da escravidão em uma posição subalternizada, tendo sua presença destinada a alimentar e servir os senhores de engenho. Conceição Evaristo, ao defender a escrevivência, aponta que sua escrita tem a intenção de “[...] rasurar essa imagem da “mãe preta” contando história. A nossa “escrevivência” conta as nossas histórias a partir das nossas perspectivas, é uma escrita que se dá colada à nossa vivência, seja particular ou coletiva, justamente para acordar

SUMÁRIO

os da casa-grande" (EVARISTO, 2017, *online*). Portanto, o conceito de escrevivência, na contramão dessa narrativa social excludente, borra o imaginário não só através da denúncia dessa construção histórica, mas também aponta caminhos emancipatórios de autonegação e afirmação, tendo a presença negra no centro discursivo, como sujeito de sua história.

UM NOVO MÉTODO PARA A CIÊNCIA DA INFORMAÇÃO

A escrevivência, inicialmente utilizada nos estudos literários, vem se dilatando para outras áreas do conhecimento, como: Ciências Sociais, Artes, Comunicação e Educação. A partir de uma pesquisa simples na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), por exemplo, é possível localizar centenas de documentos que abordam esse conceito articulado às temáticas sobre disputa de território; direito à memória; juventude negra; violência racial e de gênero, práticas pedagógicas antirracistas, saberes tradicionais e religiosos. Essa articulação conceitual demonstra que a escrevivência evaristiana se revela como um importante aporte teórico e metodológico dada a sua capacidade de proporcionar uma escrita coletiva que apresenta experiências sociais e históricas compartilhadas. A recorrência no uso da escrevivência em publicações acadêmicas aponta para a busca de conceitos teórico-metodológicos de abordagem decolonial, que possibilitam analisar o campo de pesquisa negando a hierarquização entre oralidade e escrita; corpo e mente; saber tradicional e saber científico.

Utilizando dados coletados na BDTD com um recorte localizado na construção de metodologias em Ciência da Informação, foi possível observar que a escrevivência enquanto um método de

SUMÁRIO

análise apareceu em duas dissertações de mestrado defendidas em 2022 e 2023 respectivamente: "Maniho'k: informação e memória da Casa de farinha do Povoado Terreirão" de autoria de Eva Dayane Jesus dos Santos, defendida no Programa de Pós-Graduação em Ciência da Informação, pela Universidade Federal da Bahia, e, "Maria [Beatriz] Barbosa de Souza: na gira da vida de Mãe Beata" de autoria de Karina Ceci de Sousa Holmes, defendida no Programa de Pós-Graduação em Ciência da Informação (PPGCI) do Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba (UFPB).

Para a compreensão desse método, alguns aspectos das pesquisas foram apresentados em diálogo com os elementos representativos da escrevivência evaristiana: a afirmação da pesquisadora centrada na sua experiência de vida; a memória e o ato inventivo, a escrita de si com eco na coletividade e a autodeclaração racial. A intenção foi de observar como esse método foi mobilizado na construção das pesquisas e seus desdobramentos.

Eva dos Santos (2022) identificou os elementos informacionais na Casa de Farinha do Povoado Terreirão, localizado no município de Muniz Ferreira, estado da Bahia, registrando sua origem sócio-histórica e cultural frente à apropriação oral. A abordagem da pesquisa foi de natureza qualitativa do tipo exploratória, a partir do estudo de caso e das escrevivências, as técnicas adotadas para esta pesquisa foram: entrevista, observação direta e participante; revisão de literatura e análise documental.

No texto introdutório a autora revela sua relação ancestral com o Povoado pesquisado, e indica seu afastamento do local por alguns anos gerando lacunas em sua memória. Mais adiante no texto, podemos observar o acionamento da memória através de uma imagem que retrata uma comemoração dentro da Casa de Farinha no Povoado Terreirão, foi por meio dessa fotografia que a pesquisa foi iniciada, conforme relata:

SUMÁRIO

Era 16 de janeiro de 2016, e o documento descrito foi disseminado entre o público por meio de um dispositivo móvel (smartphone) em rede social online. O acesso pela primeira vez à imagem foi feito a exatas 1 hora e 39 minutos da madrugada. Num rápido impulso, um print foi feito e guardada a imagem da cena que causara intriga e curiosidade. Aquele janeiro de 2016, certamente, materializava a informação, bem como a narrativa daquele lugar (SANTOS, 2022, p. 18).

A memória é um elemento central na escrivivência evaristiana, pois se apresenta como estratégia de sobrevivência, resgate cultural e identitário, além de propor a reescrita da história da população negra no Brasil a partir da narrativa orientada do ponto de vista de quem vivencia experiências coletivas. O acionamento da memória oportuniza a rearticulação das identidades, que são atravessadas pelo racismo e desigualdades sociais, e reivindica o ato de contar as histórias que estavam silenciadas pelos diversos mecanismos de opressão e apagamento (GOMES, 2019). Embora não haja um capítulo dedicado ao entendimento e aplicação da metodologia da escrivivência no texto dissertativo, a autora mobiliza a memória e o registro imagético para demonstrar o uso metodológico.

A segunda pesquisa analisada foi realizada por Karina Holmes (2023) e consistiu na análise de acervos pessoais de Maria Beatriz, Mãe Beata, no contexto cultural, social e religioso da cidade de João Pessoa (PB) para construir sua trajetória infomemorial. A abordagem da pesquisa foi de natureza qualitativa do tipo documental associado à entrevista narrativa. Na estrutura do texto a pesquisadora apresenta uma associação de imagens de acervo pessoal como fios condutores para o desenvolvimento dos capítulos, tendo dedicado três subdivisões para a defesa do conceito de escrivivência enquanto um método adequado para a sua pesquisa. Oralidade e musicalidade são elementos lançados pela pesquisadora a fim de marcar sua relação intrínseca com o campo de pesquisa, bem como seu posicionamento metodológico, conforme demonstra o fragmento abaixo:

SUMÁRIO

Quem pisar neste terreiro
Tem que pisar devagar
Cuidado que na porteira
Tem gente pra vigiar [...]
(Turimba/Toada para exú)

No pé da Figueira
Eu vi uma coisa linda
Pomba Gira Menina
Dançando balé
Arreia, arreia exú mulher (Turimba/
Toada para pomba gira)

Ganhei uma barraca velha
Foi a cigana que me deu
O que é meu é da cigana
O que é dela não é meu [...]
(Turimba/Toada para cigana)
(HOLMES, 2023, p. 27)

Essas características mencionadas estão em consonância com o que nos apresenta Conceição Evaristo na construção de sua narrativa: "[...] eu gosto de ler o texto em voz alta para perceber a musicalidade do meu texto, que é muito própria da linguagem oral (EVARISTO, 2017, *online*). Dentro das comunidades tradicionais e religiosas, a musicalidade se realiza através da construção coletiva, e têm o potencial de transmitir valores éticos, culturais e étnicos que orientam seus integrantes. A oralidade, por sua vez, oferece diferentes formas de elaboração conceitual que, ao se conectarem com as experiências de vida, conferem significado à produção acadêmica (GOMES, 2019).

O conceito de escrevivência vem sendo discutido no âmbito acadêmico desde 1996 a partir da publicação da dissertação de mestrado da escritora Conceição Evaristo, mas na Ciência da Informação é possível compreender seu ineditismo. Através dos fragmentos selecionados das dissertações de mestrado percebemos a incorporação da metodologia da escrevivência, evidenciada na afirmação

SUMÁRIO

do pertencimento das pesquisadoras ao campo de pesquisa; no uso de imagens de acervo pessoal, que incorrem no acionamento da memória; na proposta estilística da escrita do texto enquanto um local ritualístico, bem como a invocação da oralidade e da musicalidade - estruturas narrativas que compreendem valores de base africana. As pesquisas demonstram um anseio na construção do conhecimento a partir da realidade latino-americana com a intenção de olhar de forma particular e localizada os contextos de pesquisa para extrair deles respostas para suas problemáticas iniciais.

ESCREVIVÊNCIA NA CONTRAMÃO DA COLONIALIDADE

O entendimento de que impessoalidade e distanciamento garantem às pesquisas o tom científico, vem perdendo força e dando espaço para metodologias comprometidas em analisar o campo de pesquisa através de referenciais não hegemônicos partindo de experiências individuais e coletivas. Essas se fortalecem à medida que ganham projeção nas pesquisas realizadas, sobretudo por pesquisadoras e pesquisadores que estão interessadas na construção de metodologias próprias para compreender campos de pesquisas complexos e promotores de tensionamentos sociais no que se refere às questões de raça, sexualidades, gênero, disputa territorial, comunidades tradicionais, povos originários e religiões de matriz africana e indígena.

O positivismo no Brasil contribuiu para a manutenção de práticas coloniais, como a hierarquização de conhecimentos com a supervalorização do cientificismo em detrimento dos saberes tradicionais e o entendimento de que a racionalidade só existe se dissociada da mente e do corpo. Essa corrente enquanto modelo hegemônico, vem sofrendo várias críticas, sobretudo no que diz respeito

SUMÁRIO

à tentativa de universalização, neutralidade e busca da verdade pela objetividade. Dada a complexidade dos fenômenos sociais a serem investigados, foi possível compreender a limitação de metodologias dessa natureza e o surgimento de novas perspectivas para a busca de uma ciência mais comprometida com as transformações sociais no século XX (GOMES; DIAS, 2020).

Portanto, a utilização da escrevivência evaristiana enquanto método de pesquisa acompanha a discussão sobre os tensionamentos sociais e amplia as possibilidades metodológicas de cunho decolonial, auxiliando na observação, participação e análise das realidades sociais atravessadas pelas relações de racismo, sexismo, exclusão social e epistêmica.

Por fim, a escrevivência, concebida a partir da afirmação do “eu” promove a desobediência epistêmica (MIGNOLO, 2008), ao passo que, quem enuncia torna-se protagonista da sua história, rejeitando as narrativas e epistemologias ocidentais que tornam o sujeito um mero objeto de pesquisa. O uso dessa metodologia oportuniza a emancipação de pesquisadoras (es) negras, ao passo que sua participação no processo de construção da ciência torna-se ativa e crítica, desvelando posicionamentos que potencializam as transformações sociais, no campo de pesquisa e também no âmbito acadêmico. Possibilita questionar as estruturas que alicerçaram historicamente o conhecimento científico e promove a busca por uma contranarrativa construída a partir do olhar de dentro da experiência, com uma leitura complexa da realidade vivida e das subjetividades das relações presentes no tecido social.

O uso das metodologias decoloniais consiste no ato de reivindicar espaços de conhecimento que reconheçam e valorizem vozes e saberes diversos, buscando promover uma pluralidade de perspectivas epistemológicas e desafiar as hierarquias estabelecidas no campo do saber científico.

SUMÁRIO

COMPREENDENDO AS IMAGENS
DE CONTROLE E SUA APLICAÇÃO
COMO METODOLOGIA DE ANÁLISE
CIENTÍFICA DECOLONIAL

Imagens de controle é um conceito criado pela socióloga e feminista negra estadunidense Patricia Hill Collins (2009) para denunciar e questionar os discursos e os padrões sociais aplicados às mulheres negras. Para ela, analisar as imagens de controle revela os “contornos específicos da objetificação das mulheres negras, bem como as maneiras pelas quais as opressões de raça, gênero, sexualidade e classe se interseccionam” (Collins, 2019, p. 139). As imagens de controle são traçadas para naturalizar o racismo, o sexismo, a pobreza e outras formas de injustiça social, fazendo com que elas pareçam naturais, normais e inevitáveis na vida cotidiana.

A análise de pesquisas sobre a educação brasileira mostra que o sistema de ensino discrimina a população negra e que são observados mecanismos que influenciam negativamente nos índices de permanência e sucesso na escolarização de alunos afro-descendentes. Uma dessas formas de discriminação e preconceito na escola, é o uso de livros didáticos que naturalizam o branqueamento de seu público.

Antes de 1850, que foi quando ocorreu a abolição do tráfico de escravos, a figura e a presença negra não existia nos livros publicados mesmo tendo grande participação no cotidiano social. A abolição do tráfico, ocorrida em 1850, forçou os escritores brasileiros a voltarem sua atenção aos escravos, em particular à maneira como eram tratados. Nos textos literários desse período, os escravos eram descritos com: desgosto, piedade e de forma desumana.

SUMÁRIO

Em 1856 surge o primeiro romance abordando a temática do escravo, intitulado *O Comendador*, na obra escrita por Pinheiro Guimarães. *O Romance Til* (1872), de José de Alencar: no final da fase indianista, dentro da tendência romântica, a presença negra é manifestada como forma de contracenar com o índio. Porém, se o índio por natureza era corajoso e profundamente orgulhoso de sua independência, a presença negra era retratada segundo o estereótipo advindo da escravatura e projetado segundo uma índole humilde e resignada. A fase naturalista/realista (1881- 1883) da literatura brasileira é inaugurada com um romance intitulado. *O mulato* (1881), de Aluísio de Azevedo, em que uma personagem negra é pela primeira vez o protagonista do enredo literário. A obra denuncia o preconceito de raça e a estreiteza do horizonte que dominam o meio provinciano e que impedem Raimundo, um rapaz negro, de se casar com uma moça branca da sociedade local. O rapaz termina assassinado e a moça casa-se com um português. Para pôr a nu o preconceito, o autor descreve o rapaz como um mulato fino, educado bacharel em direito pela Universidade de Coimbra, procurando demonstrar que mesmo com todo o prestígio social, intelectual, a presença negra continua sendo alvo de discriminação (Batista; Carvalho, 2009).

A literatura infanto-juvenil brasileira, enquanto produto nacional, tem início com Monteiro Lobato. Percebe-se nas obras desse autor a continuidade da tendência da literatura para adultos: preocupação com questões nacionais, sociais, morais. Em relação à presença negra, os preconceitos e estereótipos também foram transpostos da literatura dos adultos para a literatura infantil. A Tia Anastácia, por sua vez, é caracterizada como analfabeta e chamada de "negra de estimação"; o Tio Barnabé mora fora da casa, isolado do sítio; e o Saci encarna tudo que uma criança não deve fazer, ao mexer nas coisas alheias, fazer trapalhadas, assustar as pessoas.

De maneira geral as características marcantes de pessoas negras na literatura brasileira eram e ainda se perpetuam tendo ênfase em estereótipos como a sensualidade exagerada, bestialidade,

SUMÁRIO

feiura e subserviência. A literatura pós-escravidão foi fortemente influenciada por duas correntes de pensamento em vigência na sociedade intelectual brasileira, ambas transplantadas da Europa: O Darwinismo Social e o Positivismo, de August Comte. O Darwinismo social pregava uma superioridade branca, e o positivismo dizia que os negros têm seus defeitos, mas também tem suas qualidades, podem ser melhorados (Cardoso, 2010).

No Brasil a arte desenvolvida, ligada aos cânones coloniais estéticos herdados da tradição e instituições europeias, materializam parte dessa representação, pois as imagens produzidas intencionam a confirmação de uma suposta superioridade racial da população branca, associando pessoas negras a situações degradantes.

Adiciona-se a essa perspectiva caricata que outros artistas, inclusive europeus, produziram imagens que questionavam a escravidão e valorizavam os elementos da cultura de matriz africana que sobreviviam e alteravam as características da sociedade brasileira em formação.

Debret, registra a presença negra de homens e mulheres como principal objeto de representação, tendo interesse também em compor uma iconografia do trabalho escravo no Brasil, no qual pessoas negras são retratadas como indivíduos exóticos. Suas imagens são muito reproduzidas nos livros escolares, principalmente as que retratam o castigo de pessoas negras, situações materializadas em figuras emblemáticas e traumáticas, em que há o poder de difundir e perpetuar a superioridade e o domínio sobre o negro.

Figura 1 - Uma senhora em seu lar, 1823, Debret

Fonte: ENCICLOPÉDIA Itaú Cultural de Arte e Cultura Brasileira.

Em uma de suas pinturas mais famosas, está o quadro “Uma senhora em seu lar”, de 1823. Como podemos observar na imagem acima, as pessoas negras são representadas em tamanhos proporcionalmente menores às imagens das personagens brancas e aparecem sempre em um plano inferior, sentadas no chão. A única figura negra que está de pé, parece ser uma criança e está numa posição de servidão carregando uma bandeja com um copo.

Essa construção imagética não é por acaso, e reflete apenas um contexto cruel da conformação social brasileira que foi por séculos afetada por uma ideia de superioridade branca e de exploração e animalização de corpos de pessoas negras.

SUMÁRIO

O PENSAMENTO FEMINISTA NEGRO

Angela Davis, em seu livro “Mulheres, raça e classe”, conta a triste história da mulher negra norte-americana escravizada. No decorrer de suas reflexões a autora nos apresenta como é compreendida no imaginário a figura de uma mulher branca e de uma mulher negra. Davis destaca que as mulheres negras sempre trabalharam mais que as mulheres brancas (Davis, 2016), o que explica a necessidade de um novo pensamento feminista, pois o trabalho, que foi pauta de reivindicação de movimentos feministas embranquecidos, sempre esteve presente na vida de mulheres negras desde a fundação da nação. *“As mulheres negras não eram “femininas” demais para o trabalho nas minas de carvão e nas fundições de ferro, tampouco para o corte de lenha e a abertura de valas”* (Davis, 2016, p. 22).

Davis também fala sobre a feminilidade exaltada no século XIX, onde as mulheres brancas são vistas como mães protetoras, parceiras e donas de casa, amáveis com seus maridos; e as mulheres negras eram anomalias sociais.

Após a abolição do tráfico as mulheres negras passaram a ser avaliadas em função de sua fertilidade; o que obviamente fez aumentar a violência praticada contra esses corpos.

Herbet Aptheker aponta que as discussões sobre mulheres negras escravizadas transitam entre a promiscuidade sexual ou seus pendores matriarcais. Davis observa discussões que envolvem: promiscuidade *versus* casamento; e sexo forçado *versus* sexo voluntário com homens brancos.

As linhas discursivas observadas tanto por Aptheker quanto por Davis demonstram o viés imagético criado em torno dos corpos de mulheres negras que de maneira preponderante as colocam como objeto de desejo; promíscuas, envolventes, raivosas e aptas

SUMÁRIO

para o trabalho braçal. Angela Davis também analisa sob uma ótica do que seria uma ideologia da feminilidade como um subproduto da industrialização, do modelo patriarcal e capitalista, o papel das revistas femininas e romances do século XIX e início do século XX. Nessas publicações as mulheres brancas são retratadas como princesas, mães, donas de casa e as mulheres negras vivem um contexto de total exclusão do mundo.

Patrícia Hill Collins, em seu livro “Pensamento feminista negro”, traz uma importante reflexão acerca dos estereótipos, sendo estes não apenas simples classificações, mas um instrumento de poder e controle sobre corpos historicamente subalternizados. Collins nos diz que:

Como parte de uma ideologia generalizada de dominação, as imagens estereotipadas da condição de mulher negra assumem um significado especial. Dado que a autoridade para definir valores sociais é um importante instrumento de poder. Grupos de elite no exercício do poder manipulam ideias sobre a condição da mulher negra. Para tal exploram símbolos já existentes, ou criam novos. Hazel Carbel sugere que o objetivo dos estereótipos não é refletir ou representar uma realidade, mas funcionar como um disfarce ou mistificação de relações sociais objetivas. Essas imagens de controle são traçadas para fazer com que o racismo, o sexismo, a pobreza e outras formas de injustiça social pareçam naturais, normais e inevitáveis na vida cotidiana (Collins, 2019, p. 172).

Devido ao processo de escravização, as mulheres negras receberam o status de outsider, por serem consideradas sempre o “outro”, já que o sujeito normalizado é sempre o branco e aceito socialmente é o branco. As mulheres negras também funcionam como limitadoras da margem, já que estão do lado de fora. Desta forma funcionam como ameaças, mas também são fundamentais para a sobrevivência dessa estrutura social pois explicam seus próprios limites. A partir de outra contribuição intelectual, Collins nos traz o conceito de *outsider within*,

SUMÁRIO

considerando que desde o período de escravização, essas mulheres colocadas à margem e desumanizadas, assistiram e ouviram várias cenas importantes do dia a dia das famílias brancas, já que executavam o trabalho de dentro das casas, cuidando da cozinha, alimentação, dos cuidados com as crianças de seus senhores. Esse *status* de *outsider within* tem proporcionado às mulheres afro-americanas um ponto de vista especial quanto ao *self*, à família e à sociedade.

Segundo Collins, o *status* de *outsider within* dá às mulheres algumas características especiais, como uma grande objetividade, confiabilidade nas pessoas para que se abram sobre assuntos diversos ainda que seja um estranho, e a habilidade do “estrangeiro”, em ver padrões que dificilmente podem ser percebidos por aqueles imersos na situação.

No contexto norte americano, Patrícia Hill Collins identifica que a primeira imagem de controle aplicada a mulher é a da mammy-a servicial fiel e obediente. A segunda imagem de controle é da negra má, que surge de um relatório do governo intitulado *In The Negro Family: The Case for National Action* (A família negra: em defesa de ação nacional) que apresentava a tese de que a ausência das matriarcas de suas casas para trabalhar faziam com que elas negligenciassem os cuidados com seus próprios filhos contribuindo para diversos problemas sociais. Esse modelo de mãe, era marcado também por agressividade e castração ao próprio marido, fato que explicaria o abandono desses parceiros e o fato de seguirem suas vidas como mães solo. A negra má é uma imagem de controle cruel pois culpabiliza sempre a mulher negra por qualquer adversidade que ela venha a passar.

A autodefinição e a autoavaliação das mulheres negras envolvem desafiar o processo de validação do conhecimento político que resultou em imagens estereotipadas externamente definidas da condição feminina negra. Se definir antes dos outros é a maneira que mulheres negras encontraram para se colocarem no lugar que desejam e não no lugar que a sociedade espera.

SUMÁRIO

A APLICAÇÃO DO CONCEITO DE IMAGENS DE CONTROLE COMO METODOLOGIA DE ANÁLISE CIENTÍFICA

A aplicação do conceito de imagens de controle tem sido utilizada em diferentes áreas do conhecimento, embasando análises de cunho decolonial. Em um estudo de Carvalho e Barbosa (2023), o conceito de imagens de controle foi utilizado para realizar uma análise do ensino de arte nas escolas de educação básica. De acordo com os autores:

Nesse estudo, nos apropriamos do conceito de imagens de controle para pensar e problematizar os discursos sociais, políticos e educacionais que reiteradamente aprisionam e manipulam o ensino de arte a partir de uma ideologia de dominação que menospreza o conhecimento sensível em detrimento do conhecimento lógico e racional, aquele que pode ser aferido pelas avaliações em larga escala. Partindo desta compreensão identificamos seis imagens de controle relacionadas ao ensino de arte: 1. Arte como área de conhecimento menor; 2. Alienação sobre a importância da arte na educação; 3. Desvalorização do conhecimento sensível; 4. Ensino de arte como perfumaria; 5. O reiterado discurso da polivalência; 6. O desconhecimento do que é e como se organiza o ensino de arte. Cada imagem de controle será apresentada a partir de "vinhetas do cotidiano", e articulada à produção acadêmica a respeito da arte na educação.

A partir das categorias criadas, os autores, apresentaram cenas do cotidiano que naturalizam estereótipos sobre o ensino de artes, como seu valor menor diante de outras áreas do conhecimento. A utilização dessas cenas vem para desmitificar e dialogar juntamente com um referencial teórico levantado que essa naturalização e diminuição deste campo é extremamente prejudicial para desenvolvimento do aluno prejudicando seu direito ao lúdico e prejudicando sua capacidade crítica.

SUMÁRIO

Utilizar o conceito de imagens de controle como metodologia de análise, os proporciona:

1. Categorizar melhor nosso objeto de análise a partir dos este-reótipos cristalizados socialmente;
2. Demonstrar ao leitor uma outra visão a partir de quem vivencia a situação;
3. Trazer elementos baseados em vivências e práticas cotidianas;
4. Apresentar estudos específicos relacionados ao objeto de análise que são desconhecidos pela maioria pois são invisibilizados pelo senso comum predominante.

Esse método, portanto, nos permite definir melhor um objeto e como consequência criar categorias mais consistentes, obtendo resultados confiáveis. Ainda proporciona reflexões sobre diversas ideologias, dentre elas as políticas, econômicas e sociais.

ANÁLISE EPISÓDICA: CONTRIBUIÇÕES DE GRADA KILOMBA PARA ANÁLISES SOBRE O RACISMO COTIDIANO

Advinda da psicologia, Grada Kilomba (2019) apresenta em seu livro “Memórias da Plantação” um método utilizado por ela em sua tese de doutorado para analisar episódios de racismo cotidiano. A essa abordagem, a autora denomina “análise episódica” apresentada como um caminho para analisar entrevistas realizadas pela autora, que desvelam aspectos relativos a um universo de manifestações racistas das mais diversas ordens a partir de um espaço periférico de construção do conhecimento. De acordo com a autora sua escrita reveste-se de uma política necessária ao desvelamento do racismo.

SUMÁRIO

De acordo com a autora

Não escrevo do centro, escrevo da periferia. Este é também o lugar da minha teoria, pois situo o meu discurso na minha própria realidade. O discurso das/os académicas/os negras/os surge muitas vezes como discurso lírico e teórico que transforma a linguagem do saber clássico. Um discurso que tem tanto de tão político quanto de pessoal e poético, como a escrita de Frantz Fanon ou de bell hooks. Deve ser este o principal interesse da descolonização do saber, de modo a «permitir que se produza conhecimento alternativo emancipatório», como argumenta Irmingard Staeuble, para transformar «as configurações de conhecimento e poder e abrir novos espaços de teorização e prática». Como escritoras/es e académicas/os negras/os, transformamos as configurações do conhecimento e do poder ao mesmo tempo que nos movemos entre fronteiras opressivas, entre a margem e o centro (Kilomba, 2019, p. 57).

Os episódios de racismo se constituem em análises interpretativas fundamentadas na teoria psicanalítica e pós-colonial de Franz Fanon. É com base nas proposições teóricas desse arcabouço teórico psicanalítico que Kilomba (2019) desmembra a fala das entrevistadas em análises a respeito do racismo cotidiano. São as narrativas que, advindas de corpos negros, irão proporcionar a construção de análises a partir da teoria supracitada, e a partir também da própria experiência de Grada Kilomba como mulher e pesquisadora negra.

Dessa maneira, Grada Kilomba (2019), por meio de entrevistas, examina como o racismo se manifesta no cotidiano de mulheres negras, apresentando episódios em um formato que se assemelha a contos, nos quais narra as situações enquanto as analisa à luz da teoria de Fanon. A análise episódica, segundo Grada Kilomba (2019, p. 88):

[...] descreve diferentes contextos nos quais o racismo é performado, criando uma sequência de cenas do racismo cotidiano. A composição de vários episódios revela não apenas a complexidade de experienciar o racismo - seus

SUMÁRIO

cenários diversos, atores e temas-, mas também sua presença ininterrupta na vida de um indivíduo. Essa forma de análise episódica também me permite escrever com um estilo similar à forma de contos, que, como descrito anteriormente, transgride o modo acadêmico tradicional.

Essa metodologia se insere no movimento da decolonialidade, partindo do princípio de que as subjetividades da pesquisadora importam no processo de análise da história que está sendo narrada, rompendo, assim, com a lógica tradicional de distanciamento do “objeto” de estudo. Um aspecto fundamental dessa abordagem é a utilização do formato de contos, que se configura como uma prosa científica, permitindo à autora, na qualidade de mulher negra, expressar seu ponto de vista a partir também de sua própria experiência de vida.

Grada Kilomba (2019) aponta as principais críticas que experienciou a partir da realização deste trabalho, sendo uma das mais frequentes o estranhamento por parte de outros cientistas diante de um método que foi descrito muitas vezes por eles como “não-científico” ou “muito pessoal”, uma vez que considera experiências de vida de mulheres negras e da própria pesquisadora no desenvolvimento das análises. A este respeito, discorre a autora:

Quando eles falam é científico.
Quando falamos é não científico.
universal/específico
objetivo/subjetivo
racional/emocional
imparcial/parcial
Eles têm fatos, nós temos opiniões
Eles têm conhecimento, nós temos experiências
(Kilomba, 2019, p. 52)

Neste pequeno trecho de Memórias da Plantação, Grada Kilomba (2019) critica a construção hierárquica do saber que impõe uma divisão entre o que é considerado científico e legítimo, em contraposição ao que é rotulado como conhecimento não científico e,

SUMÁRIO

portanto, sem valor. Sua ênfase encontra-se baseada na relação que caracteriza na dicotomia entre o sujeito e os outros, categorias criadas a partir da colonialidade e que se perpetua no imaginário coletivo.

A autora aponta as dicotomias que estruturam a produção de conhecimento dominante, como a oposição entre o universal e o específico, o objetivo e o subjetivo, o racional e o emocional. Além desses aspectos a obra aborda o silenciamento de pessoas negras no ambiente acadêmico, questionando “quem pode falar” nestes espaços, enfatizando que há uma “hierarquia violenta que determina quem pode falar” (Kilomba, 2019, p. 52), sendo esta hierarquia, construída sobre pilares racistas e sexistas, responsável por apagar e marginalizar sistematicamente o conhecimento produzido por pesquisadoras negras.

Ademais, a respeito do controle de “quem pode falar” Grada Kilomba (2019) apresenta o retrato de Anastácia, desenhado por Jacques Arago, um explorador francês que esteve no Brasil em uma “expedição científica” entre dezembro de 1817 e janeiro de 1818. Na ilustração, a mulher negra escravizada aparece com uma máscara que lhe impede de falar. A partir dessa imagem, Grada Kilomba (2019, p. 33-34) reflete sobre a boca como símbolo da fala, afirmando que, “no âmbito do racismo, a boca se torna o órgão da opressão por excelência, representando aquilo que as/os brancas/os querem – e precisam – controlar e, conseqüentemente, o órgão que, historicamente, tem sido severamente censurado”. A partir desta constatação, a autora reflete sobre como o próprio ato de falar emerge como um ato de insurgência, e que ela, como mulher e pesquisadora negra, não quer mais ser apenas objeto de estudo de outras pessoas, ela quer estar no centro, quer ser ouvida, e quer ser protagonista no processo de construção do conhecimento.

Estudos que se valham da Análise Episódica como alternativa metodológica para a compreensão de questões de caráter étnico-racial têm sido desenvolvidos embrionariamente em áreas

SUMÁRIO

do saber no Brasil. Se considerada a produção científica com vista a obtenção de graus acadêmicos, uma análise efetuada a partir do conjunto de registros recuperados Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), nos indica o uso pouco explorado nas teses e dissertações defendidas no Brasil, totalizando 11 registros, dentre os quais apenas 8 se constituem como métodos utilizados em estudos étnico-raciais.

Violência contra as mulheres e saúde mental: silenciamentos e invisibilidades do sofrimento de usuárias da atenção primária à saúde em Recife, de Mirela Lucena Motta (2017), problematizou os campos da atenção primária à saúde, da saúde mental e da violência, tendo como objetivo geral analisar como a violência é edificadora do sofrimento psíquico de mulheres usuárias da atenção primária à saúde em Recife. O estudo foi pautado nas epistemologias feministas em diálogo com a psicologia social, tendo como alternativa metodológica entrevistas episódicas realizadas com três mulheres na Rede de Atenção Primária no distrito sanitário IV do município de Recife. As entrevistas foram analisadas a partir dos episódios de sofrimento e de violência trazidos pelas mulheres, na perspectiva da análise episódica. Tal análise permitiu a visualização de três eixos temático-analíticos: 1) intersecções entre gênero e classe social; 2) função de cuidado e 3) saúde mental na atenção primária. Como resultados, foi possível perceber que as questões que envolvem o sofrimento psíquico das mulheres surgem na atenção primária de diversas formas, seja na reprodução da lógica, por profissionais e usuárias, de medicalização do sofrimento, na dificuldade de relatar questões entendidas pelas mulheres como mais pessoais aos profissionais, nos diferentes casos que expressam situações de opressão, com o surgimento de queixas como depressão e tentativas e/ou desejo de suicídios, ou ainda no não reconhecimento da unidade de saúde da família como instância primeira para o atendimento às situações de violência contra as mulheres. O sofrimento psíquico tem sido experimentado e percebido pelas mulheres de formas distintas, a partir das singularidades dos casos, de acordo com um leque

SUMÁRIO

amplo de violências que se articulam entre si a partir das relações sociais de gênero, classe e raça herdadas de lógicas coloniais.

A pesquisa de mestrado apresentada no ano de 2023, por, junto ao Ana Cláudia Bianconi Programa de Pós-Graduação em Educação, da Universidade de São Carlos tem o título “A gente tinha que se virar sozinha: processos educativos emergentes da andarilhagem de mulheres atletas de futsal” atribui a devida ênfase aos processos educativos que emergem da prática social da andarilhagem de mulheres atletas de futsal. A partir de registros em diários de campo de uma roda de conversa, cinco mulheres atletas de futsal que integravam o elenco da equipe principal do clube Associação Sãocharlense de Futsal – ASF manifestaram suas compreensões através de entrevistas com as participantes, as quais foram gravadas e transcritas em diários de campo. Para a análise das livres manifestações a autora se valeu da Análise Episódica, em que excertos das entrevistas foram selecionados, apresentados e discutidos, de forma a promover a reflexão conjunta em torno dos episódios associando-os, sempre que possível, a conceitos de Paulo Freire. Os resultados encontrados apontam para a evidência de processos de desumanização, a partir da opressão experienciada por mulheres atletas de futsal na andarilhagem, desvelam práticas emancipatórias e educativas de enfrentamento, de cuidado mútuo, de apoio e da constituição de identidades. Essas manifestações tem o propósito de alicerçar e contribuir para o processo de desenvolvimento da consciência crítica e como consequência de *práxis* libertadoras.

A dissertação intitulada O que estou fazendo aqui: a experiência do racismo cotidiano na trajetória universitária, de autoria de Pedro Henrique Machado (2023), se valeu da Análise Episódica para identificar o impacto do racismo no cotidiano acadêmico e profissional de pessoas negras que ingressaram no ensino superior a partir das políticas de ação afirmativa (política de Cotas Raciais). A investigação atentou para a percepção dos profissionais negros no que tange aos elementos desencadeadores de sofrimento psíquico

SUMÁRIO

catalisado pelo racismo, buscando refletir acerca das repercussões dos eventos racistas. As discussões teóricas se baseiam em diferentes formas de racismo, a presença e perpetuação de manifestações e desdobramentos da colonialidade no âmbito das atividades acadêmicas. O estudo de cunho qualitativo, desenvolvido na área da Psicologia Social e das relações raciais teve como *corpus* de análise 15 entrevistas, realizadas com base nas entrevistas narrativas biográficas de profissionais negras/negros, de três regiões do país - sul, sudeste e nordeste, que utilizaram as ações afirmativas para ingressar no ensino superior.

Estudo desenvolvido para a obtenção do grau de mestre intitulado Encontrando meu corpo: diálogos entre negritude, saúde e descolonização, de autoria de Gabriel Alves Godoi, (2023), apresentou uma ampla articulação entre os conceitos de branquitude, direitos humanos, universalismo, meritocracia, capitalismo, políticas de inimização e saúde. A compreensão desses aspectos se concretizou a partir de conceitos de negritude, determinação social da saúde e descolonização. A dissertação problematizou o acesso e o cuidado à saúde a partir de memórias e narrativas de pessoas negras sobre serviços do Sistema Único de Saúde, ao propor tensionamentos acerca do conceito de saúde com base na perspectiva da negritude, na determinação social em saúde e no pensamento negro descolonial. Como metodologia, foram realizadas interlocuções com pessoas autodeclaradas pretas ou pardas, que responderam a um questionário on-line. Inicialmente, as pessoas selecionadas foram convidadas a relatar suas experiências com o SUS através do método horizontal e circulante da conversação. As narrativas concedidas foram lidas em sua totalidade e, posteriormente, selecionadas aquelas que, por alguma circunstância, envolveram-se relevantes. Também foi utilizado como fonte de obtenção de dados o diário de campo, no qual foram anotados pensamentos, divagações, sentimentos e gestos do próprio pesquisador em meio às conversas.

SUMÁRIO

Os dados obtidos foram analisados segundo a Análise Episódica, de tal forma que foram arrolados em meio ao conteúdo das entrevistas aspectos relativos aos conceitos norteadores das discussões teóricas, quais sejam: negritude, descolonização e determinação social da saúde. A Análise Episódica permitiu através das informações expressas nas entrevistas a articulação entre sujeito e objeto, a partir da relação de amparo fundamentada no diálogo com outros os entrevistados, de tal forma que possam ser compreendidos reciprocamente a concepção de saúde a partir das experiências racializadas de ambos. Na ótica do autor “encontrar a nós mesmos significa, em grande parte, tocar naquilo que há de mais doloroso produzido pelo racismo e pelo epistemicídio”, aquilo que no fundo procuramos tanto matar, mas que nunca morre, e que, ao não morrer, nos traz força. Esse poder de vida que procura introduzir no diálogo com a saúde e sua determinação social.

Racismo, bucalidade e promoção da saúde bucal da população negra (2024) de autoria de Beatriz Carlos Correa Dulianel, apresentado para a obtenção do grau de mestre em Saúde Coletiva: Políticas e Gestão em Saúde, na área de Política, Gestão e Planejamento mapeou e analisou formas pelas quais o racismo afeta as dimensões sociais, culturais, simbólicas e psíquicas que atravessam/produzem bucalidades de pessoas negras, ao discutir como a construção da noção de bucalidades negras pode orientar as ações de atenção à saúde bucal. O método qualitativo de coleta de dados se fundamentou em entrevistas individuais abertas com seis membros do Movimento Negro Unificado de Campinas. Procedimentalmente o material das entrevistas foi transcrito, sendo adotada a Análise Episódica, criada por Grada Kilomba, compreendida nesse contexto de produção como técnica. Concomitantemente foram produzidos contos elaborados com inspiração nas entrevistas, como recurso que sustenta a liberdade de ficcionalização em um compromisso que se refere à “energia da imagem” e aos sentimentos presentes nos relatos das/os entrevistadas/os. Os participantes

SUMÁRIO

do estudo, representantes do movimento negro, acadêmica/os das áreas de saúde bucal coletiva/bucalidades e de racismo, formaram uma roda de conversa a partir de temas emergentes das entrevistas, articulados sob a perspectiva mais geral da noção de bucalidades negras, uma noção ainda não existente, para cuja construção a pesquisa intencionou a projeção de pistas iniciais.

Com base nas entrevistas foram produzidos quatro mapas que articulam pares de bucalidades emergentes na pesquisa: a bucalidade invisibilizada e a bucalidade hipervisibilizada; a bucalidade silenciada e a bucalidade entrincheirada ou belicosa; a bucalidade deserotizada e a bucalidade hipersexualizada; a bucalidade adoecida e a bucalidade em busca de uma cura. A partir desses mapas, foi possível perceber o quanto o racismo incide nas bucalidades de pessoas negras impedindo que as mesmas sejam plenas. Da mesma forma, encontramos nos próprios relatos dos entrevistados alternativas fundadas em perspectivas das africanidades que podem auxiliar no processo de cura da população negra. Intitulada Em razão de sua raça: os impactos psicossociais das desqualificações das masculinidades negras (2024), pesquisa de mestrado desenvolvida junto Programa de Pós-Graduação em Estudos da Condição Humana, da Universidade Federal de São Carlos, na linha de pesquisa Desigualdades e Diferenças no Contemporâneo, por Daniel Amâncio de Oliveira, teve por objetivo apresentar e contribuir com o diálogo acerca das masculinidades e da saúde mental da população negra, dedicando-se aos impactos psicossociais na construção da masculinidade negra na sociedade brasileira e no mundo, devido ao processo sistemático da desqualificação de suas existências.

O objetivo principal dessa dissertação foi o de investigar as consequências do racismo no processo de saúde e os possíveis sofrimentos psíquicos de homens negros, a partir de entrevistas com quatro deles. Por meio da análise episódica escolhida junto a Kilomba (2019) foram identificados os seguintes impactos psicossociais: dificuldade na construção de uma identidade positiva de si,

SUMÁRIO

interiorização dos estereótipos negativos sobre homens negros presentes na necessidade de performar um alto desempenho sexual, uma masculinidade viril, autocobrança em ser o melhor, não ter direito a errar, sentimento de preterimento e rejeição. Além disso, também encontramos relações sociais prejudicadas e desvantagens concretas devido às leituras sociais a partir de fantasias coloniais que os representam como perigosos e hipersexualizados, correspondendo a um processo de objetificação e inferiorização. Superando nossa hipótese inicial, também foi possível identificar formas diversas de enfrentamento e construção de identidades positivas capazes de subverter os estereótipos negativos e algumas opressões, bem como a descoberta de redes de apoio que fortalecem essa resistência, a partir da família de origem, religião, relacionamentos amorosos, cuidados com a saúde mental e aquilombamento. Dessa forma, apontam-se caminhos possíveis para uma abordagem e acolhimento adequado no cuidado da saúde mental de homens negros em uma perspectiva antirracista, anticolonial, antimanicomial e antimedicalizante.

No âmbito da Ciência da Informação apenas um estudo, que tratou de aspectos relacionados ao cotidiano de mulheres negras, foi localizado. A dissertação Episódios de injustiça epistêmica na vivência e prática acadêmica das pesquisadoras negras da Ciência da Informação, de autoria de Letícia Pereira de Souza (2024), no qual a autora investigou as percepções das pesquisadoras negras sobre as implicações da injustiça epistêmica no meio acadêmico. A problematização teórica, ao abordar os conceitos de epistemicídio, colonialidade, decolonialidade, movimentos epistemológicos da Ciência da Informação, injustiça epistêmica, pacto da branquitude e privilégio epistêmico; tencionou o quanto um universo de violências se tornam reincidentes em relação às pesquisadoras negras da área. A partir de uma abordagem qualitativa e exploratória se valeu de entrevistas semi-estruturadas realizadas com sete pesquisadoras negras da Ciência da Informação. Os episódios relatados pelas

SUMÁRIO

pesquisadoras foram apresentados sob pseudônimos escolhidos por elas: *Beatriz Nascimento*, *Beyoncé*, *Débora*, *Maya Angelou*, *Expedita*, *Maria Antonieta* e *Aqualtune*, demonstrando a necessidade de respeito a aspectos éticos exigidos no desenvolvimento de pesquisas dessa natureza. O método utilizado para análise das entrevistas foi a análise episódica de Grada Kilomba, permitindo que se chegassem a s análises dos episódios se dividiram em: Materialização da injustiça epistêmica na vivência e nas práticas acadêmicas das pesquisadoras; a discriminação durante a jornada acadêmica; o não-pertencimento em situações de tomada de decisões; políticas institucionais e a perpetuação da injustiça epistêmica na academia e estratégias de enfrentamento. É possível a partir das análises considerar que injustiças epistêmicas se sustentam pela deslegitimação e desacreditização do conhecimento proveniente de grupos sociais considerados subalternos, como as mulheres negras. De acordo com as análises em seus espaços de atuação, as pesquisadoras negras enfrentam diversas situações de não reconhecimento de suas contribuições epistêmicas. Ademais, as ações institucionais perpetuam tais injustiças por meio de formas de silenciamento, da falta de conscientização e da ausência de punição diante das violências por elas sofridas, além da deslegitimação de suas vozes diante de suas denúncias. Entretanto os episódios desvelam concomitantemente que as pesquisadoras continuam resistindo ao ambiente hostil que as cerca, buscando apoio entre seus pares acadêmicos que compartilham de suas dores e realidades semelhantes.

Destaques de cunho epistemológico necessitam ser trazidos à tona em razão de se constituírem em elementos que pautam a desobediência epistêmica necessária ao estabelecimento da coerência interna e externa das investigações decoloniais, razão de que as questões de cunho étnico-raciais, ao terem o racismo como elemento fundador lançam mão de explicitar contradições herdadas de lógicas coloniais opressoras. Ademais, reconhecemos que as análises das entrevistas realizadas de forma diretiva desvelam o

SUMÁRIO

rechaço e descompasso em relação a validação e reconhecimento de saberes produzidos por pessoas negras. A própria autora destaca que **em** relação ao conhecimento produzido por mulheres negras, Kilomba (2019), a partir de sua vivência, explica que:

Devido ao racismo, pessoas negras experienciam uma realidade diferente das brancas, e, portanto, questionamos, interpretamos e avaliamos essa realidade de maneira diferente. Os temas, paradigmas e metodologias utilizados para explicar tais realidades podem diferir dos temas, paradigmas e metodologias das/os dominantes. Essa “diferença”, no entanto, é distorcida do que conta como conhecimento válido. Aqui, inevitavelmente tenho de perguntar, como eu, uma mulher negra, posso produzir conhecimento em uma arena que constroi, de modo sistemático, os discursos de intelectuais negros/os como menos válidos (Kilomba, 2019, p. 54).

A questão da validade apresentada pela pensadora indica um cuidado redobrado necessita ser pensado no que tange a dimensão ontológica da pesquisadora (es) e sua relação com o fenômeno ao qual esteja se vinculando, ou seja, é necessário que o pesquisador compreenda seu *locus* de enunciação, de tal forma que a historicidade e experiências que constituíram a sua subjetividade sejam levadas em conta, de tal forma que desde a escolha do tema, tendo por referência necessidades e demandas coletivas, até os desdobramentos procedimentais tenham como referência em questão a compreensão de sua racialização.

Merece destaque outro aspecto central ao método da análise episódica que se refere a técnica, e variações dela advindas, utilizadas pela (s) pesquisadora (s) (es): a entrevista não diretiva. De acordo com Chizotti (2005) a técnica tem origem na abordagem psicoterapêutica, centrada no paciente e intenciona a obtenção de informações pautada nas falas, discursos ou narrativas das entrevistadas (o) (s), considerando que o indivíduo-fonte é possuidor dos requisitos e das competências necessárias para proferir livremente a sua experiência.

SUMÁRIO

Nesse sentido, a técnica reforça o processo interlocucional entre o pesquisador e informante, ao considerar que a força da dialogicidade esteja fundamentada em manifestações que possibilitem emergir contradições, significados, concepções, percepções e ideias situadas historicamente e compartilhadas pelas pesquisadoras (es) e os participantes dos estudos.

Nesse sentido as técnicas de entrevista se constituem em manifestações que possibilitam, a partir de um livre relato, identificando facetas que compõem o racismo cotidiano, considerando seus desdobramentos no plano imaginário, simbólico e concreto de pessoas negras. Esse desvelamento ocorre, sobretudo, a partir da identificação de violências sofridas e estratégias de (re) existência que necessitam ser constantemente articuladas. Considerando a perspectiva intelectual da autora

Embora não se possa transcrever com fidelidade toda a complexidade da vivência do racismo, a análise episódica descreve os diferentes contextos nos quais o racismo é performado, criando uma sequência de cenas do racismo cotidiano. A composição de vários episódios revela não apenas a complexidade de experienciar o racismo – seus cenários diversos, atores e temas –, mas também sua presença ininterrupta na vida de um indivíduo (KILOMBA, 2019, p. 88).

Ainda que o conteúdo das narrativas advindas das análises episódicas seja obtido individualmente, as mesmas proporcionam interpretações do quanto categorias de interpretativas obtidas *a priori* ou *a posteriori* são reincidentes e corriqueiras; ressurgimento como um relato coletivo alicerçado em práticas coloniais racistas.

Por fim deve ser enfatizado que a contribuição de Grada Kilomba (2019) se apresenta como alternativa valiosa para estudos qualitativos na Ciência da Informação que busquem analisar a intrínseca relação entre informação, etnia e raça; considerando um amplo leque de assujeitamentos fundamentados na mediação e circulação

SUMÁRIO

de informações, possibilidades não exploradas pelas pesquisadoras (es) da área. Nesse sentido, as experiências de vida de pessoas não brancas se mostram como alternativa de desvelamento de fenômenos humanos e sociais, especialmente no contexto das dinâmicas de racismo e opressão; ao adotar a Análise Episódica como método decolonial. Permite que possamos nos valer de um método reconhecendo que a constituição de subjetividades a partir das experiências pessoais são fruto de uma série de informações que incidem e são constituintes dos indivíduos a partir de sua experiência pessoal cotidiana, fonte legítima de conhecimento advindos de grupos historicamente discriminados ao dar voz às suas experiências, perspectivas e memória; desafiando as estruturas epistemológicas hegemônicas que ao longo dos tempos ignoram e desqualificam as essas narrativas.

CONSIDERAÇÕES FINAIS

As proposições apresentadas neste espaço de escrita se configuram como fundamento para a compreensão de métodos de pesquisa a serem utilizados por pesquisadores da Ciência da Informação para desvelar fenômenos informacionais relacionados às questões de etnia e raça. As reflexões buscaram apresentar a partir Grada Kilomba (2019) ao mencionar sobre a noção de epistemologia e a respeito do conhecimento produzido, bem como atribuir o devido destaque aos conceitos de outro lugar epistemológico (FANON, 2008) e desobediência epistêmica (MIGNOLO, 2008). Um outro lugar epistemológico leva em consideração que possamos nos debruçar sobre a utilização de métodos decoloniais no âmbito da Ciência da Informação, não se resumindo a o que tradicionalmente se desenvolveu ao longo da instituição do projeto da modernidade, que tiveram início no século XVI, na qual eram demarcadas dicotomias, a partir

SUMÁRIO

de aspectos disjuntivos, tais como sujeito e objeto; teoria e prática; fato e valor científico; objetividade em detrimento da subjetividade, forjando uma perspectiva fetichizada de epistemologia.

Buscamos a contribuição para elucidar a questão epistemológica Grada Kilomba (2019) ao mencionar sobre a noção de epistemologia e destacar o quanto o conhecimento produzido por pessoas negras e não brancas. A apresentação dos métodos foi precedida por compreensões relativas ao destaque a um aguçado olhar a ser assumido pelas pesquisadoras (es), atentando para o fato de que as pesquisadoras (es) segundo a vertente decolonial, devem se pautar por uma postura transgressora e vigilante. Isso porque os estudos étnico-raciais devem ser conduzidos por indivíduos pertencentes a esses grupos subalternizados, em oposição a tradição eurocêntrica se permitiu realizar: transformar indivíduos e grupos racializados em objetos de estudo da branquitude. Nesses casos uma lógica instrumental baliza a proposição das investigações e as pesquisas projetam "no outro" a possibilidade de que apenas se constituam como o observado; o que fornece os elementos para a compreensão do fenômeno científico, concebida como uma fonte meramente passiva de obtenção de dados e informações.

Em relação à escrevivência na Ciência da Informação as discussões apresentadas acenam para a necessidade de que a pertinência do ineditismo metodológico que se anuncia à área. Quando da seleção dos fragmentos selecionados nas investigações percebermos a incorporação da metodologia da escrevivência, evidenciada na afirmação do pertencimento das pesquisadoras ao campo de pesquisa; no uso de imagens de acervo pessoal, que incorrem no acionamento da memória; na proposta estilística da escrita do texto enquanto um local ritualístico, bem como a invocação da oralidade e da musicalidade - estruturas narrativas que compreendem valores de base africana. Portanto, a utilização da escrevivência evaristiana enquanto método de pesquisa acompanha a discussão sobre os tensionamentos sociais e amplia as possibilidades metodológicas

SUMÁRIO

de cunho decolonial, ampliando as formas de observação, participação e análise das realidades sociais atravessadas pelas relações de racismo, sexismo, exclusão social e epistêmica.

As imagens de controle por sua vez se consagram como método que possa ser utilizado por pesquisadoras (es) da Ciência da Informação, em virtude de que os experiências de um cotidiano de vida, materializadas na escrita das pesquisadoras é desveladora do quanto essa composição é orientada pelas pela afirmação dessas mulheres negras na sua trajetória existencial e no rechaço de questões relacionadas a estereótipos estrategicamente orientados considerando aspectos da constituição de uma subjetividade e que a partir desse posicionamento, constroi e elabora a sua própria imagem cultural, de acordo com formas próprias de narrar e analisar acontecimentos cotidianos, históricos e culturais.

A análise episódica por sua vez nos trouxe a perspectiva de revelar a partir dos sujeitos e seus lugares de fala o quanto as experiências individuais em meio ao contexto das coletividades, através de livre enunciação e da compreensão sobre ausências devam ser privilegiadas.

Os três métodos elencados trazem a baila ênfases nos indivíduos e nas mazelas cotidianas que insistem em ganhar novas roupagens desde o colonialismo. A permanência das relações de opressão advindas da composição etnia e raça, ratificam uma relação de permissividade quanto a quem fala, sobre o que fala. Nesse sentido os fenômenos informacionais são reveladores da opressão colonial perpetuada e que se amalgamam a aspectos contemporâneos, tais como o racismo epistêmico e acadêmico que prepondera na lógica das instituições de produção do saber. Ao propormos as metodologias decoloniais aqui abordadas como alternativas para os estudos étnico-raciais na Ciência da Informação que adentram a academia e se coadunam ao processo de pesquisadoras (es) negras acenamos como possível referência ao questionamento de modelos

SUMÁRIO

hegemônicos e eurocêtricos, desvelando métodos pertinentes aos seus objetos de estudo, concepções políticas e contextos de atuação. Como por exemplo Nego Bispo, que a partir da sua experiência de vida no quilombo, pensa no conhecimento sintético (eurocêntrico) e orgânico (quilombo). Então são lógicas próprias de elaboração do conhecimento que se desenvolvem fora do contexto hegemônico e que querem lançar mão de inquietações em detrimento de verdades.

REFERÊNCIAS

CHIZZOTTI, A. **Pesquisa em ciências humanas e sociais**. 7 ed. São Paulo: Cortez; 2005.

COLLINS, Patrícia Hill. **Pensamento feminista negro**: conhecimento, consciência e as políticas do empoderamento. São Paulo: Boitempo, 2019.

DOS SANTOS, Guilherme Marcondes; MARQUES, Roberto. Gramáticas do Atlântico Negro: Virgínia Bicudo e Grada Kilomba. **Revista Estudos Feministas**, [S. l.], v. 30, n. 2, 2022.

ENCICLOPÉDIA Itaú Cultural de Arte e Cultura Brasileira. São Paulo: Itaú Cultural, 2025. Disponível em: <http://enciclopedia.itaucultural.org.br/obras/111125-uma-senhora-brasileira-em-seu-lar>. Acesso em: 03 de abril de 2025. Verbete da Enciclopédia

EVARISTO, Conceição. Literatura negra: uma poética de nossa afro-brasilidade. 1996. 162 f. **Dissertação** (Mestrado em Letras) - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1996.

EVARISTO, Conceição. Conceição Evaristo: 'minha escrita é contaminada pela condição de mulher negra'. **Nexo Jornal**, São Paulo, maio 2017. Disponível em: <https://www.nexojornal.com.br/entrevista/2017/05/26/conceicao-evaristo-minha-escrita-e-contaminada-pela-condicao-de-mulher-negra>. Acesso em: 02 de março. 2025.

FANON, Franz. **Pele negra, máscaras brancas**. Salvador: Edufba, 2008.

GOMES, Elisângela; DIAS; Luciene de Oliveira. A triangulação enquanto estratégia de diálogo em pesquisa científica. **Comunicação & Sociedade**. São Bernardo do Campo, v. 42, n. 1, p. 31-51, jan. abr. 2020.

SUMÁRIO

GOMES, E. Falas insubmissas: memória e comunicação na obra da escritora Conceição Evaristo. 2019. 107 f. **Dissertação** (Mestrado em Comunicação) - Universidade Federal de Goiás, Goiânia, 2019.

HOLMES, Karina Ceci de Sousa. **Maria Beatriz Barbosa de Souza**: na gira da vida de Mãe Beata. João Pessoa, 2023. 337 f. : il. Orientação: Bernardina Maria Juvenal Freire de Oliveira. Dissertação (Mestrado) - UFPB/CCSA.

KILOMBA, Grada. **Memórias da plantação**: episódios de racismo cotidiano. Rio de Janeiro: Cobogó, 2019.

KILOMBA, Grada. **Descolonizando o conhecimento** - Uma Palestra-Performance, 2017. Disponível em: <https://www.geledes.org.br/descolonizando-o-conhecimento-uma-palestra/>. Acesso em: 3 abr. 2025.

MIGNOLO, Walter. Desobediência epistêmica: a opção descolonial e o significado de identidade em política. **Cadernos de Letras da UFF** – Dossiê: Literatura, língua e identidade, n. 34, p. 287- 324, Rio de Janeiro, 2008.

PROENÇA FILHO, Domício. A trajetória do negro na literatura brasileira. **Estud. av.**, São Paulo, v. 18, n. 50, Apr. 2004. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142004000100017. Acesso em: 10 mar. 2025.

QUIJANO, A. Colonialidad y Modernidad-razionalidad. In: H. Bonillo (Org.). **Los conquistados**. Bogotá: Tercer Mundo Ediciones; FLACSO, 1992. p. 437-449.

RODRIGUES, Dayane Silva. **Grupo como dispositivo socioeducativo-dialógico**: reflexões sobre uma intervenção com adolescentes em meio aberto. 2017. xiv, 232 f., il. Tese (Doutorado em Processos de Desenvolvimento Humano e Saúde)—Universidade de Brasília, Brasília, 2017.

SANTOS, Eva Dayane Jesus dos. **Maniho'k**: informação e memória da Casa de farinha do Povoado Terreirão. – Salvador, 2022. 169 f.: il. Dissertação (Mestrado – Programa de Pós-Graduação em Ciência da Informação) - Universidade Federal da Bahia – UFBA.

2

Jackson da Silva Medeiros

PROBLEMATIZAÇÕES E PRECAUÇÕES DE MÉTODO EM CIÊNCIA DA INFORMAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.2

SUMÁRIO

COMO CITAR:

MEDEIROS, Jackson da Silva. Problematizações e Precauções de Método em Ciência da Informação. In: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

“Quanto ao motivo que me impulsionou foi muito simples. Para alguns, espero, esse motivo poderá ser suficiente por ele mesmo. É a curiosidade – em todo caso, a única espécie de curiosidade que vale a pena ser praticada com um pouco de obstinação: não aquela que procura assimilar o que convém conhecer, mas a que permite separar-se de si mesmo. De que valeria a obstinação do saber se ele assegurasse apenas a aquisição dos conhecimentos e não, de certa maneira, e tanto quanto possível, o desca-minho daquele que conhece?” (Michel Foucault, *História da Sexualidade 2: o uso dos prazeres*)

Em meados de 2025, recebi o convite do professor Fabiano Couto Corrêa da Silva para escrever este capítulo. Consciente da performatividade das promessas (Austin, 1962¹), fiz uma intervenção de pronto: “quanto tempo tenho para isso?”. Seguiu-se 1 segundo de silêncio e uma resposta abafada por um riso: “pra ontem!... mas tem 40 dias”. Trabalho aceito, um possível norte me foi dado: um capítulo capaz de realizar uma discussão sobre metodologia e, ao mesmo

1 Utilizo a datação original das obras como gesto simbólico, metodológico e crítico frente à uma tendência dominante que, sob a alegação de utilizar “bibliografia atualizada”, reproduz uma lógica positivista e capitalista de organização do saber. Essa lógica, marcada pela substituição da densidade conceitual pela novidade editorial, opera segundo critérios de mensuração, desempenho e obsolescência típicos da racionalidade instrumental, valorizando o “recente” como se o pensamento fosse regido pelo mesmo ritmo da inovação técnica e de mercado que, pelo produtivismo bibliográfico, ignora o lugar da reflexão como trabalho demorado, que exige retorno, escuta, comparação e argumentação. A recusa dos textos fundantes não revela apenas negligência com a tradição, mas submissão à ideologia tecnocrática, industrial, mercadológica, como se os conceitos também devessem se submeter aos fluxos do discurso do capital. Reafirmar as datas originais é, portanto, resistir a esse processo raso e, simbolicamente, reconduzir o pensamento à sua espessura. Na seção “Referências”, há entre colchetes ao final de cada elemento, a explicitação do ano de publicação da obra original.

SUMÁRIO

tempo, permitir a alunos e alunas de pós-graduação conhecimento sobre a metodologia empregada nos trabalhos que executo.

Isso trouxe à tona indagações que constantemente me coloco, advindas da minha escrita e do meu modo de pensar, mas que muitas vezes, admito, evito enfrentar frontalmente como forma de tratar o conteúdo e não os caminhos que me levam a ele. Perguntas como: "que métodos utilizo?", "qual ferramental disponho para análises?", "como descrevo um modo de pensar que passa ao largo de aplicações daquilo que se convencionou chamar 'informação' dentro de uma Ciência da Informação?", entre outras. Percebi que aqui estava a possibilidade, ainda que de forma breve, de encarar e refletir sobre a posição de não determinar um método, não estabelecer procedimentos resolutores de problemas universais ou que se colocam *a priori* em relação ao objeto, mas que se guiam por um arcabouço. Com isso, não me refiro a um esquema estruturado ou a um método formalizado, mas a um modo de funcionar que encontra em Foucault uma atitude teórico-analítica que se constitui pelo permanente exercício da analítica dos regimes de verdade e das condições de emergência dos saberes e práticas e pela recusa a fundamentos transcendentais.

Esse arcabouço, entendo, está relacionado ao próprio ato de pensar como problematização na medida em que não busca aplicar regras, mas examinar condições de possibilidade dos discursos, tensionando limites do que se pode dizer, ver, fazer e pensar em determinado momento histórico. Em lugar de um caminho resolutivo, fixo, o que firmo é uma abertura que aposta no inacabamento e na instabilidade como forças produtivas do pensamento.

Este é um caminho em que a explicação é tortuosa, árida, pois não formalizar um método significa uma impossibilidade de construção positivista e reprodutiva do próprio método, indo ao encontro de processos e passos constitutivos a partir das problematizações criadas. Assim, sem a problematização, não há como estabelecer

SUMÁRIO

o método a ser seguido. A complexidade disso pelo aporte foucaultiano pode ser entendida a partir de obras reconhecidas como *Dicionário Foucault*, de Judith Revel, e *Vocabulário de Foucault*, de Edgardo Castro, onde o verbete “problematização” aparece apenas na obra da francesa e o verbete “método” ou “metodologia” não aparece em nenhuma das obras.

Este texto não é sobre uma novidade, nem sobre um conceito, tampouco sobre uma teoria ou sobre uma metodologia – essa função, na realidade, o livro aqui organizado faz com pertinência. Não pretendo neste trabalho realizar uma apresentação histórica da Ciência da Informação, sua constituição internacional e no Brasil, mesmo entendendo que as compreensões sobre a ideia de informação tenham ganhado corpo e certa evolução (?) ao longo do tempo. Também não almejo realizar uma visão, dita, paradigmática, com separações de épocas de “pensamentos” que sustentariam um viés histórico e epistemológico da informação no âmbito dessa ciência. Para esses fins, existem trabalhos muito mais competentes e que se esforçam em explicitar quadros e linhas desses aspectos.

Este trabalho é apenas fruto da indagação mencionada. A ideia aqui é primária, isto é, pensar a relação da metodologia e da informação em sua expressão de uso, uma espécie de “teoria metodológica” que busca aportes capazes de lidar com seu tratamento considerando-a uma composição discursiva. Este capítulo parte de um diagnóstico para refletir sobre os modos pelos quais a Ciência da Informação pode pensar metodologias que sejam, ao mesmo tempo, rigorosas e sensíveis à complexidade de seu objeto, reconhecendo que todo método é uma resposta teórico-prática a um problema que é tanto de conhecimento quanto de posicionamento do objeto na trama discursivo-cultural.

Esclareço, por fim, que a utilização das expressões “problematizações” e “precauções de método” no título deste texto são referências à obra de Michel Foucault, tanto como um alerta contra

SUMÁRIO

metodologias universalizantes quanto como atenção às condições históricas de formação dos discursos, um tornar instáveis objetos aparentemente dados, interrogando seus regimes de verdade. Ao incorporá-las, assumo o esforço de pensar a informação como objeto não evidente, mas como construção contingente atravessada por regimes de saber-poder que deve ser problematizada historicamente.

Com Foucault, entendo que problematização é uma atividade de raciocínio diferente daquela que assumimos como natural pelo simples fato de que os problemas não são os mesmos com as atividades de conhecimento. O conhecimento interposto hoje é uma construção do que foi estabelecido e aceito como tal anteriormente e sua proposição atual passa por elementos diferentes que o constituem, não sendo possível uma passividade de aceitação da “coisa dada”, implicando em uma incontornável análise de questões pelas quais o ser pode e deve ser pensado, indo ao encontro das práticas que formam problematizações (Foucault, 1984a).

Problematização, para Foucault (1984c, p. 236), “é o conjunto de práticas discursivas ou não discursivas que faz alguma coisa entrar no jogo do verdadeiro e do falso e o constitui como objeto para o pensamento (seja sob a forma da reflexão moral, do conhecimento científico, da análise política etc.)”. Note-se que problematização, dessa maneira, designa um movimento constitutivo de algo que passa a existir enquanto objeto a ser pensado, pois entra no regime de verdade instituído e são as práticas discursivas e não discursivas que o colocam em possibilidade de reflexão, permitindo não apenas o reconhecimento de sua existência, mas situando-o como um dispositivo-objeto de saber. Judith Revel (2007) vai nesse caminho, explicando que isso se refere a uma indagação sobre a reapresentação da forma histórica de objetos, regras ou modelos em uma dada época de um certo tipo de resposta a um certo tipo de problema.

SUMÁRIO

Isso envolve analisar o objeto para ser capaz de elaborar problemas que constituem o próprio objeto em análise, pois só a questão formulada é capaz de alicerçar a reflexão, de modo que os problemas que surgem são particulares, não universais; “supo-nhamos que os universais não existem. [...] nada de interrogar os universais utilizando como método crítico a história, mas partir da decisão da inexistência dos universais para indagar que história se pode fazer”, ensina Foucault (2004, p. 5-6). É possível enfatizar, também tomando Foucault (1984c) como âncora, que essa desconstrução não visa meramente identificar respostas diante de um impasse, mas trata de interrogar os regimes de constituição das próprias perguntas, reconstruindo os modos pelos quais certas práticas, certos discursos, foram tornados pensáveis e inteligíveis em determinados contextos históricos.

Isso age na exploração das condições em que os problemas emergem a partir de efeitos do trabalho histórico. A problematização, assim, não é um reflexo de um dado empírico ou de uma carência conceitual, mas um dispositivo de produção discursiva que constitui simultaneamente objetos, sujeitos e domínios de saber, sem a busca de fundamentos fixos, mas de estratégias, de jogos de saber-poder que tornaram certas soluções possíveis, desejáveis ou necessárias em um dado momento histórico (Foucault, 1971; 1984c).

Então é preciso estar atento às problematizações que se colocam na construção de um objeto acadêmico, objeto de saber-poder, a partir de dois pontos: o do objeto a ser analisado e o método de análise. A primeira problematização, do objeto em Ciência da Informação, está ligada à informação. Pensar a informação como prática de uso parte de uma recusa pela busca de essências universais, substituindo-a por uma análise das práticas em contextos específicos, como fenômeno situado, cuja inteligibilidade só emerge no interior de práticas sociais, discursivas e institucionais. A mobilização da ideia de informação, nesse sentido, é inseparável dos contextos que a constituem.

SUMÁRIO

A tentativa de delimitar o conceito de informação como objeto científico encontra dificuldade epistemológica fundamental, qual seja, a busca por uma definição universal, estável e ontologicamente neutra. Penso que buscar esse fundamento é um empreendimento de resultados insatisfatórios e que tal dificuldade não é casual, apontando para a necessidade de mudança de perspectiva teórica que desloque a ênfase da essência para o uso, para a prática. É neste contexto que a filosofia da linguagem de Ludwig Wittgenstein, especialmente nas *Investigações Filosóficas* (1953), oferece uma pista de compreensão conceitual capaz de iluminar condições epistemológicas da informação enquanto fenômeno social e discursivo indo ao encontro de Michel Foucault.

Nas *Investigações*, Wittgenstein recusa a ideia de que o significado das palavras possa ser encontrado em uma essência metafísica ou em uma definição fixa. Ao contrário, afirma que “o significado de uma palavra é o seu uso na linguagem” (§43). Essa formulação desloca a análise para os contextos em que as palavras adquirem sentidos, ou seja, nos diferentes jogos de linguagem nos quais os enunciados circulam e produzem efeitos. Este movimento tem consequências diretas para a compreensão da informação, pois se as palavras, os termos, os conceitos são produtos de jogos de linguagem, então a informação só pode ser compreendida a partir de suas formas de uso nos múltiplos contextos em que é produzida, transmitida, interpretada.

Ao discutir a multiplicidade de usos e o caráter não padronizado das práticas linguísticas, Wittgenstein introduz a noção de “semelhanças de família” (§67) como alternativa à busca por uma essência comum entre os conceitos, afirmando que, assim como ocorre com diferentes tipos de jogos, os conceitos não compartilham uma essência única, mas estão ligados por séries de semelhanças e de relações que emergem e desaparecem, se sobrepõem e se entrecruzam. Além disso, a noção de “formas de vida” destaca que os jogos de linguagem – por extensão, os usos da informação

SUMÁRIO

– são inseparáveis das práticas sociais, das instituições, enfim, dos contextos culturais que os sustentam, pois, “a linguagem é parte de uma atividade, ou de uma forma de vida” (§23), como ordenar, descrever, produzir, relatar, formular, testar hipóteses etc. Isso apenas reforça a necessidade de pensar a informação não como uma categoria abstrata ou como um objeto dado, mas como um fenômeno construído e situado socialmente, inscrito em práticas institucionais, tecnologias mediadoras e regimes de verdade específicos.

Como mostrado no §67, Wittgenstein busca ilustrar que uma categoria (como “jogo” ou “informação”, por exemplo) não depende de uma essência única que percorra todos os casos de uso, mas, sim, de semelhanças parciais entrelaçadas em sobreposição, criando força, uma rede de usos cujas semelhanças são mais de ordem funcional e pragmática do que ontológica. Olhando essa concepção para o campo da Ciência da Informação, é possível argumentar que “informação” não se refere a uma entidade única e invariável, mas, à articulação de múltiplos fios discursivos que sustentam o uso do conceito em diferentes contextos de pesquisa.

Essa orientação abre caminho para a problematização da metodologia de pesquisa em Ciência da Informação: se a informação é um conceito dependente de usos situados, então os métodos de pesquisa devem ser construídos de modo sensível às especificidades desses usos e às condições institucionais, tecnológicas e discursivas do objeto produzido. A adoção de uma abordagem metodológica, portanto, não pode ser feita a partir de modelos abstratos e universalizantes, mas deve emergir da análise concreta dos jogos de linguagem e das formas de vida em que a informação adquire sentido.

Segunda problematização: do método. Uma metodologia de pesquisa deve nascer como um exercício de enfrentamento teórico, epistemológico e prático diante da própria constituição do objeto a ser analisado. No caso da Ciência da Informação, da noção indefinida de informação. O estatuto ontológico e epistemológico da

SUMÁRIO

informação não se configura como um dado imediato, estável ou único, mas, ao contrário, emerge como um objeto construído por múltiplas camadas discursivas, isto é, práticas institucionais, condições de produção de saber e regimes de verdade, nos termos de Foucault. Essa condição impõe tensionamento metodológico inevitável na escolha de um método, não apenas como uma opção por um procedimento técnico ou por uma abordagem instrumental, mas, antes, como implicação de uma delimitação ontológica e epistemológica sobre o que se entende pelo objeto tratado – no caso, informação.

A partir disso, é importante reconhecer que a escolha metodológica da pesquisa não pode ser concebida como um sistema fechado de procedimentos neutros ou como uma mera aplicação de protocolos já estabelecidos, uma vez que a escolha está ancorada no próprio objeto. A informação, enquanto objeto de investigação, não pode ser tratada apenas como conteúdo passível de coleta, análise e interpretação, mas deve ser vista também como fenômeno situado, performativo e atravessado por dispositivos de poder, mediações e práticas sociais que a constituem como tal. Nessa perspectiva, a definição do método corresponde a um modo específico de produzir o objeto, de delimitá-lo e de organizá-lo.

Replicar métodos ou buscar uma metodologia universal aplicável às questões de informação ultrapassa o próprio objeto – a informação. O mais adequado parece ser o reconhecimento de que um método é sempre uma estratégia contingente, elaborada em função de um recorte analítico, de um problema de pesquisa bem definido e de uma concepção epistemológica que dê conta da variedade de sentidos e usos do objeto informação, uma vez que todo método carrega, de forma mais ou menos visível, uma teoria subjacente de seu objeto. Os próprios textos teóricos são, como Foucault argumenta, textos de esclarecimento, colocando como campo empírico a própria formulação discursiva em análise; o campo empírico é o próprio objeto de prática na constituição do objeto, uma vez que são eles próprios – os textos – os operadores dessa constituição. Na Ciência

SUMÁRIO

da Informação, onde o objeto não é dado *a priori*, mas, sim, construído a partir de complexas práticas discursivas, a escolha metodológica adquire um caráter duplamente reflexivo, pois a escolha de um método constitui simultaneamente o objeto de investigação e faz com que o uso do método deva ser estabelecido em relação a outros fenômenos ou objetos, sendo necessário utilizar instrumentos suscetíveis de análise daquilo que constitui o objeto.

Em suma, a construção de uma metodologia de pesquisa em Ciência da Informação demanda o reconhecimento da complexidade constitutiva da informação. Trata-se de complexidade ontológica, na medida em que a informação pode assumir naturezas distintas dependendo do regime discursivo em que se insere, epistemológica, onde cada abordagem teórica compreende e operacionaliza a informação de modo diferente, e pragmática, dado que o uso da informação por diferentes sujeitos e em diferentes contextos redefine seus sentidos.

Em *As palavras e as coisas* (1966), Foucault expõe como a linguagem se tornou objeto, descrevendo uma transformação fundamental na posição da linguagem no interior das formações discursivas ocidentais. Enquanto na episteme clássica a linguagem ainda operava sob a lógica da representação, como aquilo que se coloca entre o pensamento e as coisas para assegurar sua ordenação e nomeação, na episteme moderna a linguagem deixa de ser compreendida como transparente e torna-se, ela mesma, um objeto de saber. Esse deslocamento não implica uma simples mudança temática, mas uma modificação nas condições de possibilidade do próprio saber moderno, pois inaugura o tratamento da linguagem dotada de historicidade, materialidade e estrutura e não se encontra mais subordinada ao regime da representação, podendo ser analisada a partir de suas regras, leis de funcionamento, desvios, transformações.

SUMÁRIO

Entendo que um percurso metodológico honesto considera tanto as possibilidades de significação e de ressignificação que uma teoria pode propor ou sofrer quanto as mutações que o objeto de estudo pode lidar, partindo para uma análise daquilo que constitui o objeto com base no que medeia sua constituição. Isso traz coerência ao processo investigativo, não deixando que algo escape de maneira silenciosa à analítica proposta, nem esteja aquém do referencial teórico e metodológico emergente e escolhido. É o “protocolo de um exercício” (Foucault, 1984a, p. 16).

Dessa forma, utilizo o ensinamento de Foucault (1984a) sobre o ato de filosofar, buscando com diligência o trabalho do pensamento sobre o próprio pensamento, fundamentado na construção que busca não legitimar o que já se sabe. Isso não significa que, do exterior, esteja-se criando leis para outros ou querendo demonstrar uma positividade ingênua, mas, pelo contrário, entender o que pode ser mudado no próprio pensamento. Isso se coaduna com

“[...] a tarefa de evidenciar alguns elementos que possam servir para uma história da verdade. Uma história que não seria aquela do que poderia haver de verdadeiro nos conhecimentos; mas uma análise dos ‘jogos de verdade’, dos jogos entre o verdadeiro e o falso, através dos quais o ser se constitui historicamente como experiência, isto é, como podendo e devendo ser pensado” (Foucault, 1984a, p. 13).

Mas e o método efetivamente? “O senhor nunca fixou um método...”, Foucault foi perguntado. “Não tenho um método que aplicaria, do mesmo modo, a domínios diferentes.” Esse trecho, em entrevista de 1977, mostra uma recusa da fixação de um método único e transferível entre domínios, mas não representa, de forma alguma, uma negação do rigor investigativo, tampouco uma propriedade relativista do método. Indica, sim, a necessidade de atenção constante às condições de emergência dos objetos e às formas históricas de problematização.

SUMÁRIO

Foucault tem ciência da problemática e das inquiuições realizadas sobre seus trabalhos, deixando claro, em *A arqueologia do saber* (1969), que os traçados percorridos na análise de temas e teorias são de complexa explicação e explicitação, realizando uma descrição-analítica dos seus trabalhos anteriores e como o método é impactado pelo campo, pelo domínio, pelo objeto a que está ligado, como em *História da Loucura*, em *O nascimento da clínica* e em *As palavras e as coisas*: “Na análise dessas estratégias, tenho bastante dificuldade de entrar em detalhes” (Foucault, 1969, p. 77).

Como dito, em vez de um método universal aplicável *a priori*, Foucault propõe uma prática de pensamento que se dá por meio de deslocamentos, conforme as configurações específicas dos discursos, saberes e relações de poder que constituem o campo ou o objeto investigado, firmando recusa à formalização metodológica em favor de uma postura teórico-crítica situada, na qual o método se ajusta às singularidades do objeto em sua constituição histórica. Isso significa que o essencial não é encontrar um ponto de aplicação de um método que permaneça constante, mas buscar o campo, o domínio, o objeto que componha com o método a cobertura da problematização e da curiosidade imposta a isso.

O que está em jogo, portanto, é uma concepção da prática investigativa não redutível a um procedimento normativo, mas modulada por problematizações contingentes que trazem à tona certas formas de experiências. O que se convencionou chamar de “método” em Foucault é, na verdade, uma analítica arqueogenealógica² que assume, desde o início, o deslocamento e a reconfiguração como partes constitutivas da própria atividade teórica, permitindo que se utilize ou forje instrumentos durante a pesquisa.

2

Não discutirei aqui a terminologia “arqueogenealogia” ou “genealogia”, suas implicações e/ou sobreposições. A utilização do termo é apenas para a explicitação do processo que perpassa toda a construção de pensar de Foucault, cobrindo toda sua trajetória.

SUMÁRIO

Permito-me um exemplo provocativo: engana-se o pesquisador que pensa coletar dados como uma atividade trivial e neutra, deixando a análise para um momento posterior. Engana-se porque a própria coleta já é uma produção. Pense-se aqui um pequeno deslocamento à ideia de “dado” como “algo dado” em si, preexistente ao ato investigativo. Se for assim pensando, pode ser encarado como uma representação da realidade que lá está a ser simplesmente recolhido, classificado, analisado. Todavia, a perspectiva adotada aqui busca demonstrar que aquilo que se toma como “dado” é sempre resultado de práticas discursivas e dispositivos metodológicos que não apenas extraem informações do mundo, mas constituem objetos, sujeitos e regimes de verdade. “Cada sociedade tem o seu regime de verdade, sua ‘política geral’ de verdade” (Foucault, 1979, p. 12) e a pesquisa é um dispositivo inserido nesse regime, jamais neutro, sempre produtor de realidades e de sentidos.

Ao considerar que a coleta de dados é já uma produção, assume-se que não existe um campo empírico puro, externo ou independente de categorias teóricas, de instrumentos, de recortes e de interesses. A coleta pode ser compreendida como produção porque recorta, seleciona, nomeia, organiza segundo referenciais teóricos e metodológicos específicos, construindo o objeto ao mesmo tempo em que o coleta. O pesquisador que se pensa mero coletor ignora que aquilo que emerge como dado é, na realidade, efeito de operações conceituais, práticas discursivas, condições materiais e técnicas que, mais do que que refletem, performam a realidade que se propõe investigar.

Partindo dessa noção, quero me ater a realizar algumas anotações às precauções de método colocadas por Foucault no curso *Em defesa da sociedade*, mais especificamente na *Aula de 14 de janeiro de 1976*, e no texto *Foucault*, publicado sob o pseudônimo de Maurice Florence, em 1984. Não realizarei a transcrição das palavras de Foucault durante o curso ou a entrevista, pois isso pode ser facilmente acessado. Faço assim, comentários sintético-analíticos sobre as precauções.

SUMÁRIO

A primeira precaução de método apresentada por Foucault em *Em defesa da sociedade* mostra a necessidade de apreender o poder em suas formas locais e regionais, abandonando a perspectiva que o limita a codificações, possibilitando revelar sua efetividade na prática que ultrapassa as regras instituídas, consolidando-se em técnicas, instituições e instrumentos materiais, inclusive violentos. Essa abordagem evidencia que o poder se exerce em sua dimensão produtiva, em práticas cotidianas, técnicas de governo, dispositivos e instrumentos que não apenas regulamentam, mas também produzem e organizam condutas, corpos e saberes. A questão trata, portanto, de uma perspectiva que revela o caráter capilar do poder e que se infiltra no tecido social por meio de procedimentos aparentemente neutros ou técnicos, como a vigilância, a estatística, a arquitetura, abrindo espaço para sua análise como estratégia imanente às relações sociais.

A segunda precaução metodológica orienta a não analisar o poder como se fosse redutível à intenção de um soberano ou de um legislador, mas em direcionar o olhar à sua face externa, enquanto prática concreta que incide sobre um campo de aplicação, produzindo efeitos reais e transformando a realidade que atinge. Com isso, Foucault rompe com concepções clássicas do poder como posse, direito ou prerrogativa centrada no sujeito que o exerce para pensar o poder como prática que se verifica e se objetiva no próprio ato de produzir efeitos sobre condutas, corpos, discursos e saberes. Assumindo que o poder não deve ser interrogado quanto às suas intenções originárias, mas quanto aos seus efeitos, pela maneira como se implanta nos processos sociais e individuais, reorganiza-o a partir de sua ação material e simbólica. Essa precaução é articulada à sua analítica de poder enquanto rede de relações estratégicas, pois revela que o poder só pode ser conhecido a partir de sua exterioridade prática, isto é, a partir de suas consequências e modalidades de intervenção. É possível dizer que o ato de conhecer o poder significa rastrear sua eficácia e seus modos de funcionamento na materialidade dos dispositivos.

SUMÁRIO

A terceira precaução de método em *Em defesa da sociedade* constitui um dos pilares da analítica de poder foucaultiana ao romper com o modelo que concebe o poder como uma substância possuída, acumulada ou concentrada em instâncias específicas que o exercem sobre sujeitos passivos. Ao afirmar que o poder não se localiza em um ponto fixo, mas é circulante, que funciona em rede, Foucault desloca o eixo analítico para a multiplicidade de relações de força que atravessam o corpo social, configurando uma dinâmica na qual cada indivíduo é ao mesmo tempo causa e efeito. Essa perspectiva implica compreender que os sujeitos não são apenas receptores passivos, mas sujeitos posicionados em redes de poder e que inevitavelmente exercem esse poder em algum grau, ainda que em condições assimétricas e desiguais. O poder é parte das relações sociais e não um bem ou uma propriedade que possa se ter ou transferir, sendo, na realidade, um fluxo de práticas, saberes, discursos e técnicas que atravessam os sujeitos, constituindo-os ao mesmo tempo como dominados e como pontos de exercício de poder. Essa precaução metodológica pode ser vista como uma contraposição a um pensamento ou análise que define posições separadas para dominadores e dominados, evidenciando a complexidade que estrutura a sociedade moderna.

A quarta precaução metodológica formulada por Foucault marca um deslocamento radical frente às abordagens tradicionais que concebiam o poder como algo emanado de um centro soberano e difundido hierarquicamente ao conjunto do corpo social, isto é, recusa tanto a visão do poder como distribuição democrática e anárquica quanto a dedução que parte de um centro soberano para explicar seus prolongamentos. O que é proposto é, ao contrário, uma análise ascendente que parte dos mecanismos infinitesimais, locais e autônomos, com sua própria história, técnica e racionalidade, para compreender como são investidos, reorganizados ou colonizados por formas mais amplas de dominação. Trata-se de reconhecer que o poder não é apenas reflexo de estruturas macro, mas emerge de práticas discursivas locais que, ao serem apropriadas por estratégias globais, permitem a eficácia das tecnologias modernas de governo.

SUMÁRIO

A quinta precaução metodológica sublinha, em Foucault, a inseparabilidade entre as práticas de poder e a produção de saberes, desestabilizando a concepção marxista clássica que relega à ideologia um papel meramente superestrutural, como reflexo invertido ou mascaramento das relações materiais de produção. Para Foucault, o exercício do poder não atua sem uma trama de saberes que não são simples representações, mas são técnicas de observação, registro, verificação, investigação que materializam regimes de verdade e tornam corpos, condutas e populações governáveis. Assim, o que se denomina “ideologia” não deve ser apreendido apenas como ilusão ou distorção da realidade, mas como um feixe de dispositivos, como arquivos, estatísticas, diagnósticos, censos etc. que integram e que sustentam mecanismos de poder em sua eficácia. Essa articulação indissociável entre saber e poder revela que o poder moderno não é apenas repressivo, mas também produtivo, constituindo objetos, sujeitos e campos de conhecimento que retroalimentam os próprios dispositivos de dominação e de normalização. Longe de poderem ser considerados meros acessórios ideológicos, esses aparelhos de saber são elementos constitutivos das tecnologias de poder, organizando e distribuindo a verdade segundo relações estratégicas, fazendo com que toda análise do poder exija, metodologicamente, a descrição minuciosa das condições técnicas e práticas pelas quais o saber é produzido e colocado em circulação.

Sob pseudônimo, Foucault caracteriza seu trabalho como *História crítica do pensamento*, o que traz consequências à própria analítica, uma vez que considera aspectos subjetivos e objetivos da compreensão proposta. Se essa história crítica do pensamento for entendida, como anota o próprio Foucault, como uma análise das condições que formam ou modificam relações entre sujeito e objeto, é este saber que deve ser investigado. Na face subjetiva, o que deve ser o sujeito, o real e o imaginário constitutivos. Sobre o modo de objetivação, analisar as condições de como algo se torna um objeto passível e possível de conhecimento, incidindo diretamente na problematização.

SUMÁRIO

Por essa via, é possível verificar os “jogos de verdade”, isto é, as regras pelas quais coisas e sujeitos podem dizer coisas verdadeiras ou falsas. Articulação de discursos e seus efeitos, bem como isso constitui um campo de análise, indo ao encontro da maneira pela qual o sujeito se insere como objeto nesse “jogo de verdade”.

Isso permite a Foucault interrogar a existência de universais para compreender suas constituições históricas, isto é, não os tomar como dados, mas na importância de compreender o sujeito dentro de certos jogos de verdades particulares, indo além da pretensão estruturalista de imposições de fora para dentro, para, aí sim, uma compreensão da experiência, daquilo que constitui o sujeito a partir da experiência com o objeto e modifica a relação ao mesmo tempo em que modifica o campo da experiência. A investigação das práticas, aquilo que “se faz” com o objeto, parece deixar claro como são concebidas tanto como modo de agir quanto de pensar e isso será a chave de inteligibilidade para a constituição do sujeito e do objeto.

Entendo que os meios pelos quais a Ciência da Informação tem alicerçado seus estudos, em sua maioria, aplicados, muitas vezes passam ao largo de uma constituição reflexiva entre a relação do método com o objeto. Sem discussão sobre o objeto, não há discussão sobre método. Ao findar, aplica-se de forma premente. Entendo que não estabelecer esses elementos analíticos é um tipo de ataque – porque o silêncio também faz parte do ataque – à compreensão filosófica do objeto da Ciência da Informação em função de uma determinada forma de civilização industrial em constante expansão, uma adaptação ao seu século e um emparelhamento de ideias com a expansão científica, técnica e industrial do mundo em sua vertente dominante que ataca a reflexão sobre a teoria da informação com ideologia tecnocrática. A consideração comprometida de elementos da Filosofia é um modo de ajudar as ciências a sair da interminável gestação que não as colocam como “verdadeiras ciências”, mas como

SUMÁRIO

um conjunto de constante estado de técnicas. A Filosofia não faz nada além de contribuir com os meios e possibilitar olhar para as composições discursivas, principalmente pela sua materialidade, tomando as técnicas como ciência, desvelando esses caminhos³.

Suportado por Foucault e Wittgenstein, advogo em prol da surpresa, da construção, do reconhecimento de sinais que possibilitam a inventividade e a fazedura, aqui e ali, de procedimentos que sejam desenhados concomitantemente ao delinear do objeto problematizado. Compreendo que isso impõe um caráter ético ao pesquisador, qual seja, o compromisso de escutinar o próprio objeto, não tomando-o como evidente, não assumindo “ele por ele mesmo”. Toda evidência depende de regimes de produção do saber; ignorá-los é tomar como dado o que é, na verdade, construído à luz das práticas e métodos que produzem a evidência. Fora disso, toda evidência se mantém opaca, sustentando apenas a aparência de obviedade.

Conceber uma metodologia como indiferente à historicidade e às práticas que constituem o objeto equivale a reinscrever a lógica de um saber normativo que opera por meio de técnicas de normalização e exerce função disciplinar. Tal perspectiva reativa uma racionalidade estratégica de poder-saber que pretende impor inteligibilidade a partir de esquemas formais exteriores. Dessa forma, pensar uma metodologia friamente conduzida como elemento que se acopla ao objeto é o estabelecimento de um regime alheio às contingências do objeto, impondo a ele uma forma externa e rígida, um efeito de normalização na tentativa de submeter a multiplicidade do real a uma matriz unificadora e funcional.

Em contraste, situar a metodologia às condições de emergência do objeto, reconhecendo seus regimes e jogos de verdade, constitui uma prática vinculada a uma política de verdade na qual o saber é inseparável dos regimes históricos de produção e validação,

3

Desloco as palavras de Althusser em *Filosofia y ciencias humanas* (1963) para o campo da Ciência da Informação a partir da sua crítica ao ataque sofrido pela Filosofia.

SUMÁRIO

possibilitando reconhecer o conjunto de procedimentos que regulam a produção, circulação e validação dos discursos em uma dada época. Pensar a metodologia como parte de uma política de verdade é uma formulação conceitualmente rigorosa, jamais rígida.

REFERÊNCIAS

- ALTHUSSER, L. *Filosofía y ciencias humanas*. In: ALTHUSSER, L. **La soledad de Maquiavelo**: Marx, Maquiavelo, Spinoza, Lenin. Madrid: Akal, 2008. [1963]
- AUSTIN, J. L. **Quando dizer é fazer**. Porto Alegre: Artes Médicas, 1990. [1962]
- CASTRO, E. **Vocabulário de Foucault**. 2. ed. Belo Horizonte: Autêntica, 2016. [2004]
- FOUCAULT, M. **A ordem do discurso**: aula inaugural no Collège de France, pronunciada em 2 de dezembro de 1970. 24. ed. São Paulo: Loyola, 2014. [1971]
- FOUCAULT, M. **As palavras e as coisas**: uma arqueologia das ciências humanas. São Paulo: Martins Fontes, 1981. [1966]
- FOUCAULT, M. **História da sexualidade 2**: o uso dos prazeres. Rio de Janeiro: Graal, 2012. [1984a]
- FOUCAULT, M. **Microfísica do poder**. 25. ed. Rio de Janeiro: Graal, 2012. [1979]
- FOUCAULT, M. **Nascimento da biopolítica**: curso dado no Collège de France (1978-1979). 2. ed. São Paulo: Martins Fontes, 2022. [2004]
- FOUCAULT, M. *Poder e Saber*. In: FOUCAULT, M. **Ditos & Escritos IV**: Estratégia, Poder-Saber. Rio de Janeiro: Forense Universitária, 2012 [1977]
- FOUCAULT, M. *Foucault*. In: FOUCAULT, M. **Ditos & Escritos V**: Ética, Sexualidade, Política. Rio de Janeiro: Forense Universitária, 2012 [1984b]
- FOUCAULT, M. *O cuidado com a verdade*. In: FOUCAULT, M. **Ditos & Escritos V**: Ética, Sexualidade, Política. Rio de Janeiro: Forense Universitária, 2012 [1984c]
- REVEL, J. **Dicionário Foucault**. Rio de Janeiro: Forense Universitária, 2011. [2007]
- WITTGENSTEIN, L. **Investigações filosóficas**. São Paulo: Fósforo, 2022. [1953]

3

Thiago Henrique Bragato Barros

ANÁLISE DE DOMÍNIO ENQUANTO UM APORTE TEÓRICO A CIÊNCIA DA INFORMAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.3

SUMÁRIO

COMO CITAR:

Barros, Thiago Henrique Bragato. Análise de Dominio enquanto um aporte teórico à Ciência da Informação. *In*: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (orgs). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

Organização do Conhecimento (OC) é um campo disciplinar que investiga, projeta e avalia metodologias sistemáticas para representar, estruturar e tornar acessível a informação em contextos institucionais, científicos, culturais e tecnológicos diversos. Embora ancorada historicamente na ciência da informação e na biblioteconomia, OC é intrinsecamente interdisciplinar: dialoga com a linguística e a filosofia da linguagem (no tratamento do significado, da polissemia e da variação terminológica), com a lógica e a ciência da computação (na formalização de relações e na construção de modelos computacionais), com a arquivologia e a museologia (na mediação entre contextos de produção, preservação e uso), e com áreas como estudos sociais da ciência e tecnologia e comunicação (na análise dos regimes de circulação e legitimação do conhecimento). Essa transversalidade responde aos desafios de ambientes informacionais cada vez mais complexos, marcados por heterogeneidade de fontes, formatos e práticas discursivas, por escalas de produção e consumo que vão do local ao global e por requisitos de interoperabilidade semântica entre plataformas e comunidades.

Nessa perspectiva, OC busca identificar estratégias conceituais e operacionais para organizar, classificar, indexar e codificar recursos informacionais de modo que se tornem descobríveis, interpretáveis e reutilizáveis em diferentes cenários de uso. O campo

SUMÁRIO

trabalha com um repertório de artefatos — comumente referido como Sistemas de Organização do Conhecimento (SOCs) — que inclui esquemas de classificação, taxonomias, tesouros, vocabulários controlados, controle de autoridade, ontologias e modelos conceituais. Esses artefatos não apenas descrevem conteúdos: eles instauram recortes de mundo e materializam decisões epistemológicas, metodológicas e éticas, influenciando o que pode ser encontrado, como é interpretado e com que efeitos práticos. Por isso, ao lado do rigor técnico (consistência terminológica, clareza conceitual, governança de versões e alinhamentos), KO também opera uma dimensão crítica, examinando implicações de poder, vieses e exclusões que podem ser produzidos por escolhas representacionais aparentemente neutras.

Este texto apoia-se em três publicações que, em conjunto, oferecem bases teóricas, metodológicas e aplicadas para a discussão aqui proposta: um artigo na revista *Knowledge Organization*, um capítulo de livro originalmente publicado em português e um artigo em inglês na revista *Informatio*. Consideradas em conjunto, essas obras permitem articular duas camadas complementares. A primeira é a camada conceitual, que discute a natureza dos SOCS, a noção de conceito e suas relações, e o lugar da evidência empírica na construção de sistemas representacionais. A segunda é a camada metodológica, que explicita protocolos para coleta, curadoria e análise de corpora especializados; estratégias de análise quantitativa e qualitativa do discurso e do léxico técnico-científico; e procedimentos de validação e transparência capazes de sustentar inferências sólidas e reprodutíveis. Essa dupla camada favorece um entendimento de OC que não se limita à engenharia de artefatos, mas incorpora o exame das condições de produção e uso do conhecimento em domínios concretos.

Ao longo das últimas décadas, diferentes modelos teóricos e práticos foram propostos para avançar a disciplina. Entre eles, destacam-se a teoria do conceito aplicada à construção de estruturas semânticas; abordagens de análise de domínio que situam os SOCs

SUMÁRIO

em comunidades de prática, tradições teóricas e infraestruturas de comunicação científica; e vertentes de análise do discurso que examinam como classificações, índices e ontologias refletem e moldam disputas de sentido. Em termos operacionais, tais modelos se concretizam na elaboração de linguagens documentárias, na definição de hierarquias e facetas, na normalização terminológica por meio de vocabulários controlados e arquivos de autoridade, na construção de taxonomias e ontologias voltadas à interoperabilidade entre sistemas e, ainda, na explicitação de políticas de indexação e notas de escopo que guiam a aplicação consistente de termos. A centralidade da representação consistente e da recuperação sistemática permanece, mas é cada vez mais acompanhada de preocupações com a aplicabilidade das decisões, com a auditabilidade de fontes e com o desenho de mecanismos de atualização e versionamento que respondam a mudanças no estado da arte e a transformações sociotécnicas.

Esse movimento disciplinar vem sendo impulsionado por transformações tecnológicas e institucionais que ampliam tanto as possibilidades quanto as responsabilidades do campo. A incorporação de padrões e boas práticas para dados conectados e reutilizáveis, o emprego de modelos conceituais formais capazes de sustentar inferências computacionais, e a integração entre SOCS e serviços de descoberta em bibliotecas digitais, repositórios e sistemas de descrição arquivística impõem exigências de precisão, coerência e documentação que excedem a mera listagem de termos. Ao mesmo tempo, a intensificação de cenários multilíngues e multiculturais, a necessidade de mapear equivalências assimétricas entre línguas e tradições, e a atenção a comunidades historicamente marginalizadas recolocam na agenda a discussão sobre justiça representacional e governança participativa dos SOCs. Nesse contexto, decisões de modelagem — por exemplo, a delimitação de notas de aplicação, a escolha de relações hierárquicas ou associativas, a inscrição de sinônimos e variantes — devem ser encaradas como decisões epistêmicas e políticas, com impactos mensuráveis na visibilidade e na circulação de saberes.

SUMÁRIO

À luz das três publicações que fundamentam este trabalho, propõe-se, portanto, uma introdução que não apenas recapitula o estado da arte, mas também estabelece um quadro metodológico integrado. Tal quadro combina procedimentos de mensuração lexical e coligacional, úteis para identificar centralidades e padrões de uso terminológico em corpora especializados, com leitura qualitativa orientada por categorias discursivas e conceituais que permitem qualificar os resultados numéricos em contexto. A articulação entre essas camadas, por sua vez, é sustentada por protocolos explícitos de coleta, normalização e análise; por critérios de comparabilidade entre conjuntos textuais de diferentes comunidades linguísticas e institucionais; e por mecanismos de validação e transparência que viabilizam a reprodutibilidade e a crítica inter pares. Em síntese, esta introdução posiciona KO como um campo que conjuga desenho de artefatos semânticos, análise empírica de práticas discursivas e reflexão crítica sobre os efeitos de suas próprias escolhas, com o objetivo de garantir acessibilidade, reusabilidade e circulação social do conhecimento em ecossistemas informacionais cada vez mais exigentes.

FUNDAMENTOS DA ANÁLISE DO DISCURSO (AD)

A Análise do Discurso (AD) torna-se particularmente fecunda quando situada no ecossistema contemporâneo da lexicografia automatizada, em que cadeias de processamento — da extração de termos e colocações à geração assistida de definições e exemplos — são mediadas por corpora extensos, métricas quantitativas e, mais recentemente, por modelos de linguagem de larga escala. Mantém-se o núcleo teórico da AD: a linguagem como prática social e histórica e o discurso como lugar de tensão entre estrutura (regularidades, normas, convenções terminológicas, esquemas de

SUMÁRIO

metadados) e evento (irrupções, reclassificações, mudanças de rótulo e de política de uso). Essa tensão não se resolve; ela é explorada analiticamente. A novidade é que, hoje, tal tensão se materializa em pipelines computacionais que produzem entradas lexicográficas, rótulos e relações semânticas e, fazendo isso, também produzem e estabilizam sentidos. Decisões aparentemente técnicas — limiares estatísticos, modelos de extração, regras de filtragem, parametrizações e prompts — funcionam, na prática, como tomadas de posição: selecionam vozes, suprimem variantes e naturalizam hierarquias.

Na base desses pipelines, ferramentas de linguística de corpus estruturam a dimensão estrutural do trabalho. Sumários automáticos do comportamento gramatical e coligacional, relatórios de concordâncias e thesauri indutivos a partir de distribuição sustentam a identificação de colocações e perfis sintagmáticos para fins lexicográficos e terminológicos. Tais artefatos oferecem insumos auditáveis à leitura discursiva, desde que suas métricas sejam interpretadas em contexto. Em paralelo, critérios de exemplaridade para escolha de sentenças de corpus, empregados para curadoria ou geração assistida de exemplos, introduzem padrões de “típico”, “informativo” e “inteligível” que, sem vigilância metodológica, podem estreitar a representatividade social e estilística. A AD ajuda a explicitar onde essa normatividade se instala e com que efeitos.

No nível da representação e da interoperabilidade, o ecossistema técnico contemporâneo estabiliza estruturas por meio de modelos e padrões como OntoLex-Lemon (e seus módulos voltados a dados lexicográficos), a família ISO 24613 (LMF) (2019) e perfis de codificação como o TEI Lex-0, além de iniciativas colaborativas como lexemas em bases abertas. Esses arranjos permitem ligar léxicos a ontologias, documentar fontes, notas, exemplos e variantes, e tornar rastreável o ciclo de decisões. Ao mesmo tempo, fornecem pontos de observabilidade para auditorias discursivas: é possível perguntar quem definiu determinada acepção, com que corpus e justificativas, quais notas de escopo acompanham usos sensíveis e

SUMÁRIO

como se decidiu a fronteira entre categoria conceitual e rótulo lexical. Iniciativas de infraestrutura e redes de colaboração têm ampliado o alcance dessa agenda, inclusive para línguas com menos recursos, mas o ganho técnico só se converte em justiça representacional quando a governança assume explicitamente o caráter político do trabalho terminológico.

Com a entrada dos modelos de linguagem, a dimensão eventual do discurso ganha velocidade. É possível gerar definições, exemplos bilíngues e rótulos a partir de instruções, avaliar exemplos gerados com métricas automáticas e acelerar alinhamentos entre esquemas. Esses ganhos, contudo, trazem riscos conhecidos: circularidade de treinamento, reprodução de vieses, apagamentos culturais e ilusão de neutralidade. A AD oferece uma gramática de verificação que atua em várias frentes: na construção de prompts e regras de filtragem, observam-se modalizações deonticas e apagamentos de agência; na seleção de sentidos e colocações, examinam-se hierarquias de uso que invisibilizam variações diastráticas e diatópicas; na avaliação de exemplos, identificam-se critérios de exemplaridade que excluem gêneros, registros e variedades; no acoplamento entre léxico e ontologia, rastreiam-se decisões que transformam dissenso sociotécnico em univocidade formal; na governança, exigem-se trilhas de auditoria que registrem deliberações, dissensos e consequências.

Diante desse quadro, um protocolo integrado que combine AD e lexicografia automatizada pode ser sintetizado, sem a rigidez de listas, como um encadeamento contínuo. Primeiro, explicita-se o escopo com perguntas que reconhecem controvérsias, públicos afetados e riscos de exclusão. Depois, constroem-se corpora combinando grandes coleções para mensuração com amostras curadas para leitura qualitativa, assegurando comparabilidade entre línguas e domínios. Em seguida, extraem-se hipóteses estruturais a partir de frequências, palavras-chave, coligações e perfis gramaticais, sempre sob a condição de retorno às concordâncias. Modelam-se os dados

SUMÁRIO

em padrões que garantam observabilidade (fontes, notas, histórico de versões). Se houver geração com modelos de linguagem, pareia-se a produção automática a filtros estatísticos e regras conhecidas, além de validação humana especializada e documentação pública de parâmetros. Por fim, aplica-se a auditoria discursiva às definições, exemplos, rótulos e relações, e consolida-se uma governança que publique critérios, decisões e diffs, incluindo indicadores de qualidade (cobertura, diversidade de exemplos, distribuição de rótulos e variação interlinguística) para escrutínio comunitário.

Os ganhos dessa integração são práticos e epistêmicos. No plano prático, reduzem-se custos de coleta e organização, melhora-se a explicabilidade das escolhas por meio de padrões e trilhas de decisão e amplia-se a cobertura, especialmente em domínios especializados e línguas com poucos recursos. No plano epistêmico, a AD reconecta a engenharia lexicográfica às condições de produção do dizer: explicita quem fala, com quais autoridades, que memórias discursivas se reativam, que diferenças são aplainadas em nome da uniformidade e que exemplos típicos excluem variações legítimas. Essa reconexão é decisiva em cenários multilíngues e multiculturais, nos quais a busca por interoperabilidade semântica não pode ignorar conflitos de sentido, assimetrias de poder e necessidades de comunidades historicamente sub-representadas.

Há limites e precauções que precisam ser assumidos no desenho metodológico. Nenhum traço textual isolado autoriza inferências amplas sobre causalidade social; é preciso confrontar o diagnóstico com dados institucionais, históricos e de uso real. Em ambientes multilíngues, comparações devem evitar isomorfias artificiais, pois traduzir rótulos não equivale a traduzir regimes de sentido. A posição do analista é parte do problema: toda descrição é intervenção, e categorias analíticas também carregam historicidade e implicações. Ainda assim, o saldo é robusto. Integrar AD e lexicografia automatizada qualifica a leitura de artefatos de organização, melhora a aplicabilidade das decisões, fortalece a governança terminológica

SUMÁRIO

e oferece instrumentos para redesenhar recursos lexicais e terminológicos mais auditáveis, inclusivos e cientificamente responsáveis, sem abdicar do rigor técnico nem da eficiência computacional. Em última instância, reconhecer que organizar conhecimento é também organizar relações de poder, memórias e expectativas de futuro permite que a tensão entre estrutura e evento seja operada com consciência crítica, transformando pipelines em espaços de reflexão e melhoria contínua.

ANÁLISE DO DISCURSO PARA A ORGANIZAÇÃO DO CONHECIMENTO

No contexto francófono, a Análise do Discurso (AD) nasce, no final dos anos 1960, do entrelaçamento de debates filosóficos, científicos, políticos e sociais, tendo Michel Pêcheux como figura central. Em diálogo com a filosofia de Louis Althusser, Pêcheux propõe uma teoria da materialidade do discurso: o discurso é simultaneamente material histórico-linguístico e material ideológico, isto é, produz sentido porque se inscreve em formações ideológicas que interpela(m) sujeitos e estabiliza(m) provisoriamente significados. Nesse quadro, a ideologia não é mero “véu” sobre uma realidade econômica, mas uma força produtiva com operações próprias; e a linguagem deixa de ser um sistema abstrato e fechado para ser apreendida como prática social situada. Daí a noção, cara a Pêcheux, de uma “semântica do discurso” que, em vez de reduzir o sentido à combinatória formal, investiga as condições de produção do dizer: posições enunciativas, memórias discursivas, pré-construídos, interdiscursos que atravessam os textos.

A AD opera teoricamente a tensão entre estrutura e evento. Por estrutura entendem-se regularidades, convenções terminológicas, normas de uso, políticas institucionais e rotinas que conferem

SUMÁRIO

previsibilidade à produção de sentido; por evento, as irrupções que deslocam tais regularidades — reclassificações, renomeações, mudanças de política, controvérsias, disputas definicionais. Essa tensão não se resolve, mas se torna o próprio objeto analítico: é no vaivém entre estabilização e perturbação que se evidenciam os mecanismos pelos quais sentidos se consolidam, circulam, entram em conflito e se transformam. Diferenciando-se do estruturalismo saussuriano, que privilegia o sistema, e de quadros formalistas que isolam a língua de suas condições sociais, a AD toma o texto como acontecimento, o sujeito como posição historicamente constituída e o sentido como efeito de relações de força.

Essa perspectiva sustenta uma implicação central para a Organização do Conhecimento (OC): vocabulários controlados, taxonomias, classificações, ontologias, arquivos de autoridade e esquemas de metadados não apenas descrevem objetos, mas instauram recortes de mundo e materializam decisões epistêmicas e políticas. A pretensa neutralidade de instrumentos que definem termos preferidos, estabelecem relações hierárquicas e associativas, fixam notas de escopo e orientam políticas de indexação precisa revelar-se como produto de formações discursivas que determinam o que pode ser dito, como e por quem. Em consequência, a AD fornece operadores para auditar SOCS e práticas de representação: modalizações deônticas e epistêmicas que prescrevem usos ("deve", "é proibido", "considera-se"), apagamentos de agência por passivas e nominalizações que naturalizam processos ("normalização", "conformidade"), intertextualidades que invocam autoridades normativas, estratégias de equivalência que achatam diferenças culturais sob rótulos unificadores e genealogias de termos que evidenciam disputas históricas por nomeação.

A ligação com a Análise de Domínio (ADo) reforça o alcance dessa leitura. A ADo descreve comunidades, instituições, gêneros e infraestruturas de comunicação; a AD interroga "como se diz" dentro desses arranjos. Trianguladas, permitem relacionar decisões de

SUMÁRIO

modelagem (por exemplo, a hierarquia entre categorias, a introdução de facetas, a redação de notas de escopo) com as condições histórico-institucionais que as tornaram plausíveis e com seus efeitos na circulação e na recuperação da informação. Assim, a AD não substitui o rigor técnico da OC; ela o adensa, reintroduzindo a historicidade e a ideologia como dimensões incontornáveis do desenho e da governança de SOCS.

Esse argumento ganha atualidade no cenário de lexicografia automatizada e de pipelines computacionais que hoje sustentam a organização do conhecimento. Ferramentas de linguística de corpus, extração terminológica e sumários coligacionais aceleram a identificação de padrões de uso e de perfis gramaticais; modelos conceituais e padrões de representação (como os que viabilizam a ligação entre léxicos e ontologias) tornam rastreáveis definições, exemplos, rótulos de uso e histórico de versões; técnicas de geração assistida, por sua vez, produzem definições, exemplos e alinhamentos em ritmo inédito. Contudo, em cada etapa desses pipelines — seleção de corpora, limiares estatísticos, filtros, parametrizações, prompts, critérios de exemplaridade — há escolhas que funcionam como tomadas de posição: selecionam vozes, invisibilizam variantes diastráticas ou diatópicas, naturalizam hierarquias, exportam consensos locais como universais. A AD oferece justamente a gramática de verificação para tornar tais escolhas observáveis e discutíveis, recolocando em primeiro plano a pergunta “quem fala, com que autoridade, para quais públicos e com quais consequências?”

Integrar a AD ao cotidiano da OC significa, portanto, adotar procedimentos analíticos que percorram do indício textual ao regime de verdade. No plano do texto, examinam-se marcas enunciativas, escolhas definicionais, notas de escopo, exemplos, rótulos de uso e decisões tradutórias; no plano do corpus e da infraestrutura, interrogam-se as origens dos dados, os vieses de cobertura, a distribuição de gêneros e registros, e a documentação de versões e “diffs” que torna auditável a evolução dos recursos. No plano institucional,

SUMÁRIO

observam-se as instâncias de governança, os ritos de validação e as formas de participação das comunidades afetadas. O objetivo não é paralisar a tomada de decisão, mas elevá-la a um patamar de aplicabilidade em que dissensos sejam documentados, alternativas sejam consideradas e riscos de exclusão sejam mitigados por notas, avisos de uso e estratégias de atualização.

Se conhecimento é campo de disputa, traduzir tal evidência em protocolos de OC implica reconhecer limites e responsabilidades. Em ambientes multilíngues e multiculturais, traduzir rótulos não é traduzir regimes de sentido; a comparação entre línguas e tradições deve evitar isomorfias artificiais e admitir assimetrias. Em sistemas automatizados, métricas e modelos não dispensam leitura qualitativa: concordâncias e amostragens críticas são necessárias para impedir que frequência se confunda com legitimidade. E na interface com usuários e comunidades, a transparência sobre critérios, fontes e justificativas é condição para que a circulação de saberes seja mais justa sem abdicar do rigor técnico.

Em síntese, a AD, especialmente na linhagem pecheutiana, oferece à OC um repertório conceitual e metodológico capaz de operar a tensão entre estrutura e evento como motor de análise e de projeto. Ao tornar visíveis as condições discursivas que sustentam classificações, vocabulários e ontologias, a AD amplia a capacidade da OC de explicitar pressupostos, documentar escolhas, corrigir assimetrias e desenhar sistemas mais auditáveis, inclusivos e responsivos às transformações sociais e tecnológicas. Isso vale tanto para SOCS "clássicos" quanto para ecossistemas de lexicografia automatizada, em que a velocidade de produção não pode dispensar a responsabilidade discursiva sobre aquilo que, afinal, se institui como "conhecimento".

SUMÁRIO

COMPILADO DE ANÁLISES JÁ EFETUADAS:
A INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION

A seguir, desenvolvo um paralelo aprofundado, de escopo metodológico e aplicado, entre três vertentes analíticas — Análise de Domínio (ADo), Análise do Discurso (AD) e Análise de Corpus — tomando como base os estudos comparativos sobre ISKO-Brasil/NASKO e a investigação sobre a ISKO-UK (Barros *et al.*, 2024 e Daros, 2024). O objetivo é explicitar como essas vertentes se articulam, em sequência e em iteração, para produzir leitura robusta de comunidades discursivas e, a partir daí, orientar a construção e a avaliação de Sistemas de Organização do Conhecimento (SOCs).

A ADo fornece a moldura sociocognitiva e histórico-institucional do domínio; a AD explicita as condições enunciativas e ideológicas que atravessam conceitos, termos e políticas de representação; a Análise de Corpus aporta mensuração empírica e replicável (frequências, palavras-chave, perfis coligacionais e gramaticais, contrastes) e obriga o retorno sistemático aos contextos de uso via concordâncias. Nos estudos da ISKO, essa tríade é operacionalizada com corpus comparável selecionado por enunciados-pivô, processamento no Sketch Engine e leitura qualitativa orientada, dando lastro à ADo e à AD em decisões metodológicas verificáveis.

Delimitação do domínio e do problema (pelo prisma da ADo). A ADo, na linhagem de Hjørland, oferece um repertório de abordagens para situar objetos, atores, infraestruturas de comunicação científica e regimes de legitimidade do domínio. Nos trabalhos aqui mobilizados, destacam-se as abordagens 6 (estudos históricos de estruturas e serviços de informação), 8 (estudos epistemológicos e críticos), 10 (estudos de estruturas e instituições de comunicação científica) e 13 (inserção da análise do discurso como contribuição metodológica), combinadas para dar conta tanto do “onde e quem fala” quanto do “como e com que efeitos” se fala no campo.

SUMÁRIO

No estudo sobre a ISKO-UK, essa moldura é explicitada também com apoio da tipologia de Tennis (2003) (Áreas de Modulação e Graus de Especialização), para nomear e dimensionar o domínio, definir escopo, alcance e propósito da análise e, assim, justificar seleção documental (artigos completos e resumos expandidos) e horizonte temporal (conferências de 2009–2023). Essa etapa torna visíveis inclusões e exclusões e posiciona a ISKO-UK como especialização no interior da ISKO internacional.

Constituição de corpus e compatibilização entre comunidades (ADo + Corpus). O estudo ISKO-Brasil/NASKO estabelece critérios comparáveis por enunciados-pivô, constrói dois corpora dimensionados (532.496 palavras em português; 583.674 em inglês) e explicita a filtragem de ruídos antes da extração de listas de frequência. O recorte brasileiro privilegia dimensões aplicadas e social/cultural/políticas; no NASKO, consolida-se uma amostra de 132 documentos (127 artigos e 5 resumos estendidos). Essa padronização permite uma leitura cruzada dos focos comunitários.

A ISKO-UK adota princípio idêntico de constituição do corpus por enunciados-pivô, reforçando que a construção do material é parte do método (e não etapa neutra). Isso garante que a amostra reflita o “discurso de dentro” da comunidade, nas suas práticas e prioridades históricas, e habilita a comparação com ISKO-Brasil/NASKO sem impor isomorfias artificiais.

Ferramental de corpus e papel dos lemas (Corpus + retorno qualitativo). Em ambos os estudos, o Sketch Engine é empregado com quatro módulos: Wordlist (frequência), Keywords (extração terminológica), Word Sketch (comportamento coligacional e gramatical) e Word Sketch Difference (contraste de coligações), com pesquisas realizadas por lemas (lexemas) para mitigar viés morfológico. Como se trata de ferramenta técnica, os dados retornam obrigatoriamente às Concordances (KWIC) para interpretação em contexto — ponto crítico para evitar reificação de métricas.

SUMÁRIO

Essa arquitetura quanti-quali aparece também na ISKO-UK, onde o uso de KWIC evidencia centralidades temáticas e aproxima objetivos institucionais do capítulo (promoção de pesquisa e desenvolvimento de SOCs). A Wordlist local mostra grande proeminência de "knowledge", "information", "system", "ontology", "classification" e "indexing", sinalizando um viés tecnológico/temático coerente com a tradição britânica do tratamento temático da informação.

Diferenças linguísticas e ajuste de estratégias. O estudo ISKO-Brasil/NASKO é explícito ao registrar que, por operarem em línguas distintas (português brasileiro e inglês americano, com variações), as comunidades exibem comportamentos gramaticais diferentes e, por isso, a mesma estratégia de busca não pôde ser aplicada uniformemente. Esse reconhecimento metodológico evita comparações enviesadas e sustenta inferências com parcimônia.

Inserção da AD como camada interpretativa (estrutura-evento, ideologia, enunciação). A AD, na linhagem francesa de Pêcheux, lê a produção de sentido como efeito de formações ideológicas. No capítulo teórico-metodológico, essa chave é mobilizada para deslocar a OC de uma engenharia pretensamente neutra para um exame das condições de produção do dizer: políticas terminológicas, notas de escopo, hierarquias e equivalências passam a ser decisões epistêmicas e políticas. Aplicada aos corpora da ISKO, a AD orienta a procura por marcas textuais (modalizações deônticas/epistêmicas, apagamentos de agência via passivas/nominalizações, apelos a autoridades normativas, traduções que aplainam dissensos).

Essa lente é incorporada como abordagem 13 na ADo: "análise do discurso para compreender o discurso institucionalizado no domínio". Na prática, ela funciona como auditoria de justiça representacional acoplada à mensuração de corpus e à cartografia institucional da ADo.

SUMÁRIO

Convergências e especificidades entre capítulos (achados comparados). O levantamento ISKO-Brasil/NASKO aponta convergências estruturais — centralidade de ADo, SOCs, Teoria do Conceito e Sistemas de Classificação —, mas evidencia diferenças de entendimento conceitual e de ancoragem epistemológica. No Brasil, “indexação” emerge como termo-chave central na Wordlist e o recorte privilegia dimensões aplicadas e social/cultural/políticas; o NASKO compõe uma série longa com ênfases próprias e estabilidade de eventos bienais. Essas distinções ajudam a explicar diferenças na “gramaticalização” das práticas conceituais em cada comunidade.

Na ISKO-UK, o quadro léxico aponta forte presença de “ontology”, “classification” e “indexing”, e um entrelaçamento entre “system”, “search”, “retrieval” e “semantic”, sugerindo prioridade a arranjos e serviços informacionais, com ênfase tecnológica e de automação. Por inferência, isso aproxima o capítulo de pautas de interoperabilidade e desenho de artefatos digitais, ao passo que mantém o eixo representacional (indexação) em posição de destaque.

Tradução metodológica em decisões de projeto para SOCs. A passagem do diagnóstico para o desenho/avaliação de SOCs segue um encadeamento replicável:

a) Seleção e normalização de termos preferidos e variantes. A análise de frequência, dispersão e coligações fornece limites operacionais para preferir rótulos, mapear quasi-sinonímias e distinguir homonímias. A leitura por concordâncias impede que a alta frequência de termos paratextuais (cabeçalhos/rodapés) leve a ruídos interpretativos, e amarra cada escolha a exemplos contextualizados.

b) Definição de escopos e notas de aplicação. A AD ajuda a formular notas que deem visibilidade a exceções, usos sensíveis e controvérsias, explicitando limites e situações de não aplicação. Na ISKO-Brasil/NASKO, o próprio desenho conceitual de KOS como “categoria-guarda-chuva” orienta o detalhamento de relações

SUMÁRIO

semânticas e o esclarecimento de sobreposições funcionais entre ontologias, tesouros e taxonomias (por inferência a partir da apresentação dos KOS como saídas centrais do campo).

c) Relações hierárquicas e associativas. Os perfis coligacionais e gramaticais (Word Sketch/WSD) permitem identificar pares prototípicos verbo-objeto, modificadores e contextos definicionais, sustentando BT/NT/RT com evidência empírica. Nas três comunidades, observam-se proximidades funcionais entre ontologias, tesouros e taxonomias que recomendam menos “oposição” e mais integração pragmática em arquiteturas de informação. Essa diretriz, fortalecida pela leitura de corpus e pela ADo, reduz arbitrariedades e favorece interoperabilidade.

d) Governança terminológica e trilhas de auditoria. ADo e AD indicam a necessidade de registrar quem decide, com base em que corpus, em que datas e com que justificativas. Em ecossistemas multilíngues, a equivalência interlinguística deve ser tratada como hipótese a ser provada por corpus e por comunidade (evitando traduções que aplainem dissensos). O estudo ISKO-UK reforça esse ponto ao vincular escolhas a tipologias documentais e a objetivos institucionais de promoção de SOC.

e) Indicadores para auditoria discursiva. Inspirados nas categorias da AD, é possível acompanhar proporções de modalizações deônticas, incidência de passivas/nominalizações, densidade de remissões normativas e diversidade de exemplos. Esses indicadores não substituem o juízo crítico, mas apontam zonas de risco (hipernormatização, apagamentos de agência, assimetrias de representação). O capítulo teórico-metodológico, ao reintroduzir ideologia/historicidade no núcleo da OC, dá a justificativa conceitual para tal auditoria.

Integração iterativa entre ADo, AD e Corpus (ciclo de validação). O método praticado nos estudos ISKO opera como ciclo:

SUMÁRIO

- mapear a comunidade (ADo), nomeando o domínio e suas subáreas;
- construir corpus alinhados a esse mapeamento, com documentação de critérios;
- mensurar léxico e perfis gramaticais/coligacionais (Corpus);
- qualificar os achados por leitura enunciativa e ideológica (AD);
- traduzir resultados em decisões projetuais para SOC (termos, escopos, relações, notas, governança);
- revalidar em subcorpora, períodos ou eventos diferentes, para testar estabilidade e sensibilidade das escolhas.

Esse ciclo está descrito no artigo ISKO-Brasil/NASKO (integração quanti-quali, uso de Concordance) e reaparece, com ênfase de ADo descritiva, na análise ISKO-UK.

Efeitos práticos sobre comunidades discursivas e seus SOC. Em ISKO-Brasil, a centralidade de “indexação” sugere priorizar políticas de indexação e notas de aplicação detalhadas, com atenção à cadeia arquivística e a práticas de descrição que aparecem no discurso da comunidade. Em NASKO, a ênfase histórica do simpósio, sua longevidade e estabilidade de ocorrência implicam consolidar glosários e mapeamentos com horizonte temporal amplo, aproveitando a massa de eventos bienais. Na ISKO-UK, o destaque de “ontology”, “classification”, “indexing” e termos ligados à recuperação reforça a necessidade de alinhamentos entre artefatos (tesauro-ontologia-taxonomia) e entre estes e serviços digitais, com governança transparente de versões. Essas inferências seguem dos perfis léxicos, dos objetivos institucionais e da leitura ADo/AD das três comunidades.

Limites, precauções e reprodutibilidade. Há riscos metodológicos que os estudos reconhecem e mitigam: ruídos paratextuais nas listas de frequência; vieses morfossintáticos entre línguas; assimetrias

SUMÁRIO

de recorte temporal; tentação de equiparar alta frequência a centralidade conceitual. A resposta envolve limpeza de dados, uso de lemas, retorno às concordâncias, documentação de parâmetros e justificativas para recortes e amostragens — todos passos descritos no artigo ISKO-Brasil/NASKO e coerentes com a ADo descritiva na ISKO-UK.

O paralelo metodológico-aplicado evidencia que: (i) a ADo estrutura a pergunta e situa o corpus; (ii) a Análise de Corpus oferece evidências e limites operacionais; (iii) a AD reintroduz ideologia e enunciação como dimensões inescapáveis nas decisões representacionais; (iv) a construção de SOC's deve ser vista como tradução de um diagnóstico discursivo de domínio em políticas terminológicas, relações semânticas e mecanismos de governança auditáveis. Em ISKO-Brasil e NASKO, essa tríade produz convergências fortes nos fundamentos e diferenças situadas na maneira de “gramaticalizar” práticas e prioridades; em ISKO-UK, reafirma-se a necessidade de articular artefatos semânticos a serviços e infraestruturas digitais, preservando a legibilidade das decisões. O quadro integrado oferece, assim, um caminho replicável para análise de comunidades discursivas e para o desenho de SOC's mais explicáveis, inclusivos e responsivos.

Em contexto bilíngue, a equivalência entre estratégias de busca e extração precisa ser calibrada às especificidades morfosintáticas e derivacionais de cada língua: no português, a variação flexional e a produtividade sufixal (-ção, -ismo, -idade) demandam lematização rigorosa e listas de stopwords disciplinares mais extensas, enquanto no inglês a alta incidência de compostos nominais (“knowledge organization system”) e a cabeça final dos sintagmas nominais recomendam tratamento sistemático de bigramas/trigramas e normalização de hífen e composições.

Para evitar isomorfias artificiais e vieses comparativos, a compatibilização metodológica deve incluir mapeamentos interlinguísticos de termos e variantes, validação por concordâncias

SUMÁRIO

paralelas e análises de sensibilidade (com e sem lemas, diferentes listas de exclusão e parâmetros de coligação), explicitando os limites de transferibilidade dos achados entre comunidades e períodos.

Assim, a “equivalência” entre corpora e procedimentos deve ser tratada não como identidade, mas como compatibilidade operacional demonstrada por testes de sensibilidade, mapeamentos inter-linguísticos e validação sistemática por concordâncias. Em termos aplicados, as decisões de modelagem em SOCs (termos preferidos, notas de escopo, relações) precisam vir acompanhadas de documentação metodológica explícita — parâmetros, versões e justificativas — para garantir reprodutibilidade, auditabilidade e transferência responsável entre comunidades.

CONSIDERAÇÕES FINAIS

Retomando a discussão desenvolvida neste trabalho, e tomando como caso aplicado a análise dos capítulos da ISKO (ISKO-Brasil, NASKO e ISKO-UK), consolida-se a tese de que sistemas de organização do conhecimento devem ser concebidos e avaliados como artefatos sociotécnicos. O encadeamento metodológico adotado — Análise de Domínio (ADo) para situar o campo e suas instituições, Análise de Corpus para mensurar padrões léxico-coligacionais em materiais representativos, e Análise do Discurso (AD) para qualificar ideologicamente os resultados — mostrou-se adequado para ler comunidades discursivas “por dentro” de seus próprios enunciados, e “por fora” de suas infraestruturas e regimes de legitimidade.

Do ponto de vista dos achados comparados, os três capítulos convergem em fundamentos e vocabulário disciplinar (centralidade de SOCs, teoria do conceito, classificação e ADo), mas diferem nas ênfases temáticas e sociotécnicas. Em ISKO-Brasil, “indexação” e

SUMÁRIO

práticas de descrição (inclusive arquivísticas) ocupam papel destacado e reorientam prioridades de modelagem e notas de aplicação. Em NASKO, a estabilidade institucional do fórum e a tradição de discussão conceitual sustentam arranjos classificatórios e glossários com horizonte temporal amplo. Na ISKO-UK, ontologias, taxonomias e integração com serviços de recuperação apontam para um viés tecnológico consistente (automatização, IA, web semântica), com repercussões diretas sobre alinhamentos e interoperabilidade. Essas diferenças, longe de fragilizarem a comparação, justificam que a construção de SOCs seja situada e calibrada às comunidades efetivamente servidas por eles.

Metodologicamente, a estratégia de constituição de corpora por enunciados-pivô e o uso combinado de Wordlist, Keywords, Word Sketch/Word Sketch Difference e Concordances deram escala e transparência às inferências. A leitura discursiva reintroduziu historicidade, posição enunciativa e formação ideológica, impedindo que estatísticas sejam tomadas como sinônimos de consenso. Em ambiente bilíngue e multicêntrico, a “equivalência” entre corpora e procedimentos não pode ser tratada como identidade: exige compatibilização demonstrada por lematização apropriada, tratamento sistemático de compostos, mapeamentos interlinguísticos de termos e variantes, validação por concordâncias paralelas e testes de sensibilidade (parâmetros, janelas de coligação, listas de exclusão). Esse cuidado reduz isomorfias artificiais e melhora a transferibilidade dos resultados entre capítulos.

No plano aplicado, o pipeline ADo–Corpus–AD se traduziu em orientações concretas para projeto e governança de SOCs. Seleção de termos preferidos e variantes deve combinar métricas (frequência, dispersão, produtividade morfológica, perfis coligacionais) e amostras de contexto; notas de escopo precisam explicitar usos sensíveis, fronteiras conceituais e assimetrias de tradução; relações BT/NT/RT devem ser justificadas por evidência empírica e racional discursivo, reconhecendo a complementaridade funcional entre ontologias,

SUMÁRIO

tesauros e taxonomias; e a governança terminológica tem de publicar critérios, fontes, trilhas de versão e justificativas, incluindo mecanismos de participação das comunidades afetadas. Em capítulos com forte ênfase tecnológica, como a ISKO-UK, esse conjunto de práticas encontra terreno imediato de aplicação na interoperabilidade entre serviços de descoberta, vocabulários e ontologias.

A incorporação de práticas contemporâneas de lexicografia/terminografia assistidas por computador — inclusive quando apoiadas por modelos de linguagem — amplia as possibilidades e as responsabilidades. Ganhos de escala e velocidade só são aceitáveis quando acompanhados de documentação pública de parâmetros, validação humana especializada, indicadores de diversidade/representatividade dos exemplos e auditorias discursivas que acompanhem modalizações, apagamentos de agência e decisões de equivalência. Em termos de transparência e prestação de contas, padrões abertos de representação lexical e conceitual e trilhas de mudanças tornam observáveis os pontos de decisão que, de outro modo, permaneceriam opacos.

Há limites a reconhecer. Assimetrias de qualidade e tamanho dos corpora, diferenças morfossintáticas entre línguas, ruídos paratextuais e recortes temporais heterogêneos recomendam prudência, análises de sensibilidade e explicitação de incertezas. Nenhum traço textual isolado autoriza conclusões totalizantes; é indispensável triangular dados discursivos com informações institucionais, históricas e de uso real. Ainda assim, o saldo do caso ISKO é claro: operar, de modo integrado, ADo, análise de corpus e AD aumenta a reprodutibilidade, a auditabilidade e a capacidade de revisão dos SOCs, sem sacrificar o rigor técnico.

Como agenda prospectiva, três frentes se impõem. Primeiro, replicar o protocolo em outros capítulos/idiomas da ISKO e em domínios temáticos diversos, para testar limites de comparabilidade e construir mapas de variação intercapítulos. Segundo, consolidar

SUMÁRIO

indicadores de auditoria discursiva (modalizações, agência, densidade de remissões normativas, diversidade de exemplos) acoplados a métricas informacionais (cobertura, precisão, interoperabilidade), permitindo acompanhar deriva conceitual e impacto sobre recuperação. Terceiro, institucionalizar processos de governança com participação qualificada, registros públicos de dissensos e ciclos regulares de atualização — condição para que transferências entre comunidades sejam responsáveis e documentadas.

A equivalência metodológica entre comunidades deve ser tratada como compatibilidade operacional demonstrada, não como identidade presumida; e, em termos aplicados, cada decisão de modelagem (termos, notas, relações) precisa vir acompanhada de documentação explícita — parâmetros, versões, justificativas — para garantir reprodutibilidade, auditabilidade e transferência responsável entre capítulos da ISKO. Organizar conhecimento continua sendo organizar relações, memórias e prioridades; o mérito do percurso aqui adotado é oferecer instrumentos robustos para fazê-lo de modo explicável, situado e eticamente responsável.

REFERÊNCIAS

BARROS, Thiago Henrique Bragato; LIED, Maria Amália Cassol; ROCHA, Waleska da Silva; MOURA, Kamila de Andrade. Knowledge organization systems (KOS) in the context of ISKO: a domain analysis of the Brazilian and North American chapters. **Knowledge Organization**, v. 51, n. 8, p. 582–599, 2024.

BARROS, Thiago Henrique Bragato. A análise do discurso de Michel Pêcheux e a organização do conhecimento: possibilidades teórico-metodológicas. *In*: AMORIM, Igor Soares; SALES, Rodrigo de (Orgs.). **Ensaio em organização do conhecimento**. Florianópolis: Editora UDESC, 2021. p. 61–78.

BARROS, Thiago Henrique Bragato. Knowledge organization and discourse analysis. **Informatio**, Montevideo, v. 30, n. 1, e208, 2025.

SUMÁRIO

BOOIJ, G. **The Grammar of Words:** An Introduction to Linguistic Morphology. Oxford: Oxford University Press, 2005.

DERRIDA, J. **Mal d'archive:** une impression freudienne. Paris: Galilée, 1995.

DOSSE, F. **História do estruturalismo:** o campo do signo (1945–1966). São Paulo: Ed. UNESP, 1993.

DAROS, J. Í. **Tendências discursivas em organização do conhecimento:** uma análise de domínio do capítulo do Reino Unido da International Society for Knowledge Organization (2009–2023). 2024. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) — Universidade Federal do Rio Grande do Sul, Porto Alegre.

FROHMANN, B. **The power of images:** a discourse analysis of the cognitive viewpoint. *Journal of Documentation*, v. 48, n. 4, p. 365–386, 1992.

FROHMANN, B. **Discourse analysis as a research method in library and information science.** *Library and Information Science Research*, v. 16, n. 2, p. 119–138, 1994.

FROHMANN, B. **Discourse and documentation:** some implications for the information field. *Journal of Documentation*, v. 57, n. 6, p. 678–691, 2001.

FROHMANN, B. **Documentation redux:** prolegomenon to (another) philosophy of information. *Library Trends*, v. 52, n. 3, p. 387–407, 2004.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain analysis. **Journal of the American Society for Information Science**, v. 46, n. 6, p. 400–425, 1995.

HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, v. 35, n. 2–3, p. 86–101, 2008.

HJØRLAND, B. Domain Analysis. **Knowledge Organization**, v. 44, n. 6, p. 436–464, 2017.

HODGE, G. **Systems of Knowledge Organization for Digital Libraries:** Beyond Traditional Authority Files. Washington, DC: Digital Library Federation, 2000.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 24613-1:2019 — **Language resource management** — Lexical markup framework (LMF) — Part 1: core model. Geneva: ISO, 2019.

KILGARRIFF, A. *et al.* **The Sketch Engine:** Ten years on. *Lexicography*, v. 1, n. 1, p. 7–36, 2014.

SUMÁRIO

- KILGARRIFF, A. *et al.* GDEX: **Automatically finding good dictionary examples in a corpus**. In: Proceedings of the 13th EURALEX International Congress. Barcelona: IULA, 2008. p. 425–432.
- LÓPEZ-HUERTAS, M. J. **Domain analysis for interdisciplinary knowledge domains**. Knowledge Organization, v. 42, n. 8, p. 570–583, 2015.
- ONTOLex-Lemon Community Group. **Ontology-Lexicon Model for Ontologies** (OntoLex-Lemon). [S.l.]: W3C, 2016.
- ONTOLex-Lemon Community Group. **Lexicography Module (lexicog) for OntoLex-Lemon**. [S.l.]: W3C, 2020.
- ORLANDI, E. P. **Análise de discurso: princípios & procedimentos**. Campinas: Pontes, 2007.
- ORLANDI, E. P. **Discurso e texto: formação e circulação dos sentidos**. Campinas: Pontes, 2009.
- PÊCHEUX, M. Análise automática do discurso. In: GADET, F.; HAK, T. (Orgs.). **Por uma análise automática do discurso: uma introdução à obra de Michel Pêcheux**. 5. ed. Campinas: Editora da UNICAMP, 2014 [1969]. p. 61–162.
- PÊCHEUX, M. Língua, linguagens, discurso. In: ORLANDI, E. P. (Org.). **Análise de discurso: Michel Pêcheux** – textos selecionados por Eni Puccinelli Orlandi. 4. ed. Campinas: Pontes, 2015 [1971].
- PÊCHEUX, M.; FUCHS, C. A propósito da análise automática do discurso: atualização e perspectivas. In: GADET, F.; HAK, T. (Orgs.). **Por uma análise automática do discurso: uma introdução à obra de Michel Pêcheux**. 5. ed. Campinas: Editora da UNICAMP, 2014 [1975]. p. 163–252.
- PÊCHEUX, M. **Semântica e discurso: uma crítica à afirmação do óbvio**. 5. ed. Campinas: Editora da UNICAMP, 2014 [1975].
- SMIRAGLIA, R. P. **Domain Analysis for Knowledge Organization: Tools for Ontology Extraction**. Waltham, MA: Chandos/Elsevier, 2015.
- TEI Consortium; BAŃSKI, P.; ROMARY, L.; TASOVAC, T. TEI **Lex-O Guidelines for Lexicography**. [S.l.]: TEI Consortium, 2019.
- ZENG, M. L. **Knowledge Organization Systems** (KOS). Knowledge Organization, v. 35, n. 2–3, p. 160–182, 2008.



Eixo 2

**ABORDAGENS
MÉTRICAS
E AVALIATIVAS**

Metodologias quantitativas e avaliativas dedicadas a mensurar a produção, o impacto e a dinâmica da informação, incluindo ferramentas como a psicometria, a análise de patentes e os diversos estudos métricos sobre a produção científica.

4

*Vanessa Inácio de Souza
Jussara Borges*

PSICOMETRIA: METODOLOGIA PARA AVALIAÇÃO EM CIÊNCIA DA INFORMAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.4

SUMÁRIO

COMO CITAR:

SOUZA, Vanessa Inácio de; BORGES, Jussara. Psicometria: metodologia para avaliação em Ciência da Informação. *In*: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (orgs). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

A sensação de competência (ou seja, o quanto a pessoa se sente capaz) é algo que não pode ser medido diretamente, como a altura ou o peso. Por isso, precisamos de outros métodos para entender e avaliar essa sensação.

A avaliação é uma importante ferramenta de gestão em programas de desenvolvimento de competências. Neste trabalho, o objeto de avaliação são as competências em informação e em comunicação (competências infocomunicacionais). A competência em informação diz respeito à capacidade de lidar com conteúdos, desde sua busca e seleção, passando pela avaliação e organização, até a sua produção e distribuição. Já a competência em comunicação refere-se à capacidade de relacionar-se com pessoas e, com elas, trabalhar e produzir colaborativamente, envolvendo, assim, escuta ativa, negociação de sentidos e empatia. Tratam-se de competências fundamentais no contexto contemporâneo, quando as pessoas estão sobrecarregadas de conteúdos e precisam discernir a informação útil, relevante e pertinente, além de tornar o ambiente digital saudável e colaborativo. Assim, várias iniciativas de promoção de competências infocomunicacionais vêm emergindo em escolas, universidades, empresas e organizações.

SUMÁRIO

A avaliação pode auxiliar essas iniciativas a definir o foco a ser tomado em relação a conteúdos e didática, por exemplo, entre outras funções. Embora a avaliação possa ser feita de diversas formas, é necessário que os instrumentos utilizados para tal fim sejam confiáveis, isto é, que avaliem de fato as competências e não outro aspecto qualquer. Por exemplo, perguntar a um estudante o que é uma fonte confiável mede apenas seu conhecimento conceitual sobre o tema; já pedir que ele analise uma notícia e justifique se ela é confiável ou não, avalia de fato a competência. A distinção é importante, pois instrumentos mal elaborados podem gerar resultados que não refletem as competências reais dos estudantes.

A Psicometria é a área que usa modelos e métodos matemáticos para medir dados coletados em pesquisas psicológicas (Jones; Thissen, 2007). Em resumo, a Psicometria é o ramo da Psicologia que cria testes para medir fenômenos psicológicos. Ela surgiu no final do século XIX, quando psicólogos europeus e americanos começaram a usar a Matemática para entender as diferenças entre as pessoas (Pasquali, 2010). Pode-se afirmar que o positivismo e o empirismo influenciaram, à época, a tentativa de compreender o comportamento humano por meio da aplicação da matemática.

Vale explicar que o termo medir, embora controverso, ainda é utilizado na literatura desta área, mas sua concepção se aplica ao uso de testes. Já o termo avaliação abarca as interpretações de testes à luz do contexto, ou seja, a medição é apenas um extrato da avaliação.

A Psicometria desenvolveu-se ao longo dos anos, tanto em estudos matemáticos quanto nos procedimentos para a criação de testes. Atualmente existem testes que medem inteligência, competências, personalidade, emoções, transtornos etc. Esses testes, de modo geral, são ferramentas que ajudam a compreender o comportamento humano. Os resultados obtidos por meio deles podem influenciar diretamente a vida das pessoas, pois servem, por exemplo, de base para a definição de políticas educacionais ou na seleção

SUMÁRIO

de candidatos para empregos. Por isso, é importante garantir o rigor no processo metodológico na elaboração dos testes.

No caso específico de programas de competência em informação promovidos em nível universitário, testes psicométricos podem ajudar estudantes a se autoavaliarem e a identificar onde precisam melhorar, especialmente no que diz respeito ao grau de percepção que possuem sobre as próprias competências, isto é, sua percepção de autoeficácia. O resultado disso pode representar um aumento no desempenho acadêmico.

Em outras palavras, tais testes ativam a autoavaliação dos respondentes, isto é, funcionam como um gatilho metacognitivo fazendo com que eles tomem consciência sobre questões da sua aprendizagem e, conseqüentemente, fazendo-os buscar melhorias em relação às competências de busca, uso e comunicação da informação.

Neste capítulo a Psicometria será abordada sob o ponto de vista de uma metodologia científica válida para a Ciência da Informação, bem como será trazido um breve relato sobre a construção de um teste de avaliação de competência em informação¹.

VALIDADE DOS TESTES NA PSICOMETRIA

As duas teorias mais conhecidas em Psicometria são a Teoria Clássica dos Testes (TCT) e a Teoria de Resposta ao Item (TRI). Ambas se dedicam à construção de testes, mas utilizam metodologias diferentes. A TCT, baseada nos estudos de Charles Edward Spearman, psicólogo inglês conhecido por seu trabalho na área de

¹ Este capítulo é uma adaptação baseada em: SOUZA, V. I. de. *Avaliação da percepção de competências infocomunicacionais (APCI): elaboração e busca de evidências de validade da escala*. 2022. Dissertação (Mestrado em Ciência da Informação) Porto Alegre, 2022. Disponível em: <https://lume.ufrgs.br/handle/10183/249769>. Acesso em: 26 abr. 2025.

SUMÁRIO

Estatística (Sartes; Souza-Formigoni, 2013) analisa o desempenho geral do respondente em testes, ao contrário da TRI, que extrai significados das respostas dadas a cada item de um teste. O principal critério da TCT é a validade. Segundo a American Educational Research Association, a American Psychological Association e o National Council on Measurement in Education (2014), validade significa o quanto as evidências e a teoria justificam as interpretações dos resultados dos testes. Em outras palavras, o teste mede de fato aquilo a que se propõe medir? Que evidências podem ser coletadas para demonstrar a precisão do teste? Por exemplo, um teste para medir a depressão em um paciente, está de fato medindo esse construto ou está medindo construtos subjacentes, como a melancolia? Percebe-se, nesse caso, que a definição dos construtos é fundamental para identificar suas características, diferenças e dimensões na hora de construir um teste.

Mas afinal, quais são as evidências para determinar a validade de um teste? Pode-se reunir diferentes tipos de evidências, como o conteúdo do teste, a maneira como as pessoas respondem, a estrutura interna do teste, a relação com outras variáveis e as consequências do uso do teste (American Educational Research Association; American Psychological Association; National Council on Measurement in Education, 2014). Essas evidências formam um conjunto e validam as interpretações das respostas. De certo modo, elas são confirmatórias e dão mais segurança para as inferências obtidas.

Essas fontes de evidência podem iluminar diferentes aspectos de validade, mas não representam tipos distintos de validade. A validade é um conceito unitário. É o grau em que todas as evidências acumuladas apoiam a interpretação pretendida das pontuações dos testes para o uso proposto (American Educational Research Association; American Psychological Association; National Council on Measurement in Education, 2014, p. 13, tradução nossa).

SUMÁRIO

Embora não seja obrigatório reunir todas as fontes de evidências, múltiplas formas de validação fortalecem o processo de construção desses instrumentos e aumentam a confiança no teste. Esse conjunto de evidências dá suporte teórico e empírico para afirmar que o teste mede, de fato, o que se propõe a medir, o que é crucial para garantir a precisão e a adequação das interpretações dos resultados obtidos.

Em Psicometria, a abordagem holística da validade é reconhecida como uma prática robusta, conforme estabelecido por organizações como a American Psychological Association (APA), a American Educational Research Association e o National Council on Measurement in Education (2014), além de outros estudos relevantes (Pacico *et al.*, 2015; Urbina, 2014). A seguir, serão apresentadas brevemente as principais fontes de validade.

FONTES DE EVIDÊNCIA DE VALIDADE

A validade de conteúdo é a primeira que deve ser considerada. Ela analisa como o tema é abordado, se as perguntas do teste fazem sentido, se elas estão relacionadas ao que se quer medir, se há linguagem incompreensível para o público e se há perguntas irrelevantes. De certa forma, o processo para validar o conteúdo traz as diretrizes para a construção do teste e é fundamental para garantir a precisão (American Educational Research Association; American Psychological Association; National Council on Measurement in Education, 2014). As questões do teste são elaboradas a partir da literatura científica sobre o construto, mas também podem ser complementadas com a experiência prática do(s) profissional(ais) que elaboram o teste.

SUMÁRIO

Em Psicologia, construto é uma característica psicológica, habilidade, ou traço abstrato que não pode ser diretamente observado, mas que se presume existir com base em evidências teóricas e empíricas (Gamst *et al.*, 2017). Em outras palavras, é um conceito como Inteligência ou Competência e, geralmente, mas não necessariamente, pode ser definido na literatura científica. O método para buscar a validade de conteúdo pode ser resumido em três etapas: verificação do conteúdo por especialistas, análise semântica e aplicação piloto do teste com o público-alvo².

A segunda evidência é a validade do processo de resposta, que examina como as pessoas respondem ao teste, ajudando a identificar falhas ou interpretações erradas das perguntas (American Educational Research Association; American Psychological Association; National Council on Measurement in Education, 2014). Pacico (2015) exemplifica afirmando que para obter essa evidência pode-se perguntar ao participante qual foi o raciocínio utilizado para responder aos itens. Isso permite ajustes para tornar o teste mais preciso.

A estrutura interna do teste também precisa ser avaliada, verificando se as perguntas estão bem relacionadas ao que se pretende medir e se estão classificadas nas dimensões, isto é, aspectos, fatores ou componentes do construto, corretas. Aqui, métodos estatísticos, como a Análise Fatorial, são usados para confirmar essa relação e categorizar os itens para fins de pontuação, garantindo resultados consistentes e coerentes (Gamst *et al.*, 2017). Para entender de forma simples: uma dimensão agrupa itens de um teste que estão fortemente relacionadas entre si porque medem algo em comum. Esses itens podem ter alguma relação com perguntas de outros grupos (ou dimensões), mas essa relação é mais fraca, então elas permanecem agrupadas de forma distinta. A Análise Fatorial

SUMÁRIO

consegue ver quais dimensões refletem certos comportamentos. Essas dimensões podem então ser consideradas os fatores causadores do comportamento. Caso se alinhem com o que a teoria previa, isso indica que o teste tem boa estrutura interna e validade de construto (Pacico *et al.*, 2015).

A relação com outras variáveis, em linguagem simples, nada mais é do que uma comparação entre os resultados de diferentes instrumentos que medem o mesmo construto ou similares. Por exemplo, um teste de esperança deve ter correlação positiva com o teste de otimismo. Por outro lado, a correlação deve ser negativa ao se comparar com um teste de depressão. Evidências discriminantes indicam que o teste não deve se correlacionar com características irrelevantes, como velocidade de processamento em um teste de esperança (Pacico *et al.*, 2015).

Por fim, as consequências da testagem também constituem evidência de validade, que podem ser tanto positivas quanto negativas, impactando a vida das pessoas. Decisões tomadas a partir dos resultados de um teste precisam ser analisadas em termos de suas implicações sociais e culturais (Urbina, 2014). Por exemplo, quais os efeitos éticos do teste: ele gera injustiças, discriminação ou preconceitos? Favorece sistematicamente um grupo em detrimento de outro (como gênero ou etnia)? Isso pode levar a consequências negativas e questionar a validade do teste.

Embora existam muitos métodos para coletar essas evidências e tentar minimizar erros nos testes psicológicos, a medição psicológica nunca será totalmente exata. Mesmo assim, os erros podem ser reduzidos por meio de cálculos, como o de erro padrão, que aproxima o resultado da realidade (Hauck Filho; Zanon, 2015; Pasquali, 2009). Um dos métodos mais conhecidos na TCT para avaliar a consistência interna de um teste é o Alfa de Cronbach, que indica o grau de precisão das respostas (Gamst *et al.*, 2017; Maroco; Garcia-Marques, 2006; Pasquali, 2009).

SUMÁRIO

A Teoria de Resposta ao Item (TRI), ao contrário da TCT, preocupa-se com a probabilidade que um sujeito apresenta de responder corretamente cada item de um teste e o que isso significa (Pasquali, 2009). Em testes de seleção é possível inferir em quais habilidades ou conhecimentos o candidato é mais forte ou se ele/ela está “chutando” as respostas. Assim, a TRI foca na qualidade individual dos itens e na construção de **itens válidos**. Com isso, é possível montar bancos de itens que podem ser usados em testes de aptidão ou psicológicos. A TRI não invalida a TCT, mas é considerada um avanço e uma resposta a muitas limitações da TCT (Craig, 2017; Karen B. Goldfinger, 2019). De qualquer forma, a metodologia para a construção de testes independe do modelo a ser utilizado para as análises dos dados resultantes da aplicação de testes.

Em 2020, as autoras deste capítulo deram início aos estudos para a construção do instrumento *Avaliação de Percepção de Competências Infocomunicacionais* (APCI), a partir de pesquisa conduzida no âmbito do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (UFRGS). A proposta foi aplicar os princípios da Psicometria para a avaliação da competência em informação e comunicação. A seção seguinte aborda a construção do instrumento, a busca por evidências de validade de conteúdo e o processo de construção dos itens.

CONCEITUAÇÃO DO CONSTRUTO E OPERACIONALIZAÇÃO DE ITENS

Segundo Borsa e Seize (2017), construir um instrumento de medição envolve duas etapas principais: teórica e empírica. A fase teórica é dividida em três partes: a conceituação do construto, a elaboração dos itens e a análise desses itens. A fase empírica não

SUMÁRIO

é o foco deste capítulo, mas a título de informação envolve análises estatísticas a partir das demais fontes de evidências.

A conceituação, na fase teórica, envolve definir o fenômeno que se deseja medir, abordando suas características e propriedades que, posteriormente, serão transformadas em itens mensuráveis (Pasquali, 2010). Nesta etapa, também se determina as dimensões do construto, que são, em uma analogia, como as subdivisões do fenômeno. Juntas, as dimensões formam o todo. Um exemplo seria a competência em informação, que precisaria ser conceituada, delimitada e dividida em dimensões como Busca, Uso, Avaliação e Comunicação da informação. As dimensões podem ser baseadas em modelos como o Framework for Information Literacy da American Library Association ou ainda em outras dimensões, como aquelas relacionadas às atitudes ligadas à competência em informação.

Na elaboração dos itens, os pesquisadores descrevem as ações e comportamentos que refletem o construto (Borsa; Seize, 2017). Eles podem utilizar modelos de competência já existentes, como o já citado Framework, o Big 6 de Mike Eisenberg e Bob Berkowitz, o Information Search Process (ISP) de Carol Kuhlthau, entre outros, como base. Além disso, é recomendável consultar instrumentos semelhantes para aproveitar itens que já foram validados, adaptando-os conforme necessário (Lopes; Pinto, 2013; Pett; Lackey; Sullivan, 2003).

Esse processo é chamado de operacionalização de itens ou definição operacional. De acordo com Pett, Lackey e Sullivan (2003, p. 6), "operacionalizar um construto é defini-lo de forma que ele possa ser medido ou identificado". Ou seja, operacionalizar significa fazer a ponte entre algo abstrato e a realidade, utilizando manifestações/ações observáveis do construto para representar o que se pretende investigar. Essas manifestações são baseadas tanto em pesquisas anteriores e modelos quanto em observações empíricas.

SUMÁRIO

Uma série de cuidados deve ser tomada na elaboração de cada item. Pacico (2015, p. 61) afirma que os itens devem conter uma pergunta por vez, isto é, devem evitar composições (geralmente com a conjunção “e”), como “finalizo minhas tarefas no prazo e inicio outras sem dificuldade, [pois] o respondente pode [...] finalizar tarefas, mas fora do prazo, o que dificulta a escolha de uma única resposta adequada.” Outro cuidado diz respeito à adaptação da linguagem ao público do teste. É recomendável evitar estrangeirismos, itens com negação (pois podem confundir o leitor), itens que demonstram desejos (ao invés de ações concretas), pouca variação na forma dos itens (para não entediar o respondente), gírias (que podem afetar a credibilidade) e palavras de frequência e intensidade como “sempre, nunca, muitíssimo” etc. Estas são apenas algumas das recomendações da literatura. Pasquali (1999) cita e explica doze regras para a construção dos itens, por isso sua leitura é recomendada.

Na prática, os itens de um teste são as perguntas ou afirmações que o compõem. Existem várias maneiras de fazer essas perguntas. Por exemplo, uma pergunta pode simular uma situação:

Imagine que você recebeu uma crítica inesperada sobre um trabalho em que investiu bastante tempo. O que você faria?

Alternativas possíveis:

- a. *Ficaria desmotivado.*
- b. *Sentiria frustração, mas buscaria entender os pontos mencionados para ajustar o trabalho.*

Outra forma comum são afirmações que o participante deve avaliar, indicando frequência, intensidade ou seu grau de concordância ou discordância, como em:

Eu me sinto nervoso ou inquieto em situações sociais.

SUMÁRIO

Alternativas:

(1) *Nunca*, (2) *Raramente*, (3) *Às vezes*, (4) *Frequentemente*, (5) *Sempre*

A escolha da melhor forma de fazer essas perguntas depende do objetivo do estudo e de uma série de fatores, como as vantagens e desvantagens de cada tipo, o tempo necessário para responder, a clareza das perguntas, o tipo de instrumento utilizado e as exigências estatísticas. Recomenda-se a leitura de Sampieri (2014) para um aprofundamento sobre os tipos de instrumentos e suas vantagens e desvantagens.

ANÁLISE DE JUÍZES, ANÁLISE SEMÂNTICA E ESTUDO PILOTO

Com a prévia do instrumento construído, dá-se início às análises de juízes, análise semântica e por fim o estudo piloto. A análise dos itens é realizada por especialistas, chamados de juízes, que avaliam a qualidade dos itens com base em sua experiência e conhecimento do tema. A escolha desses juízes deve ser criteriosa, considerando sua formação e publicações na área, e idealmente envolve entre três e dez especialistas (Borsa; Seize, 2017). Eles avaliam os itens classificando-os em termos de qualidade e pertinência, e a maioria dos autores sugere um índice de concordância de pelo menos 80% para que o item seja aceito (Pasquali, 2010). O Coeficiente de Validade de Conteúdo (CVC) que “[...] mede a proporção de especialistas que estão em concordância sobre os itens” (Borsa; Seize, 2017, p. 23), pode ser usado nesta etapa. Itens com valor abaixo de oito devem ser excluídos ou modificados. Alguns autores recomendam o uso de grupos focais nessa fase (Chang *et al.*, 2012), mas não há consenso. O importante é que os itens que obtiverem notas baixas sejam revisados ou eliminados.

SUMÁRIO

Logo após a análise de juízes e subsequentes alterações nos itens, inicia-se a análise semântica, etapa de testagem da versão preliminar do instrumento com o público-alvo, visando avaliar o entendimento dos itens. Esse processo pode ser feito individualmente, com um roteiro de entrevista, ou coletivamente, como em grupos focais. Pasquali (2017) recomenda que os participantes leiam os itens e expliquem com suas próprias palavras o que entenderam, idealmente em pequenos grupos de 3 a 4 pessoas. Já Pett, Lackey e Sullivan (2003) recomendam grupos de até 10 pessoas. As observações feitas pelos participantes devem ser documentadas a fim de fazer ajustes posteriores, com foco na clareza e pertinência. O processo pode ser repetido até que não surjam mais problemas de compreensão. Em alguns casos, é feito um estudo piloto para reforçar a validade do conteúdo.

O estudo piloto busca garantir que os participantes compreendam os itens e as instruções do teste, simulando uma aplicação real sem a presença do pesquisador. Os participantes podem deixar seus comentários no próprio formulário do instrumento ou dar *feedback* posterior de outra forma. O objetivo é verificar a aplicabilidade do instrumento em condições próximas das reais (Selau; Silva; Bandeira, 2020). Este processo pode ser repetido várias vezes até que os itens estejam adequados (Gurung, 2015).

A consulta à literatura deve ser constante em todo o processo. Por fim, todos esses procedimentos teóricos – desde a conceituação do construto até o estudo piloto – fornecem evidências de validade do conteúdo do instrumento. O quadro 1 resume todo o procedimento:

SUMÁRIO

Quadro 1 – Resumo do procedimento de construção

Definição do construto	Procedimento
Definição do construto	Buscar as características, propriedades e delimitação do construto na literatura
Elaboração de itens iniciais	Categorizar os comportamentos e transformá-los em itens
Consulta a especialistas	Verificar o CVC dos itens
Revisão linguística e cultural dos itens (público-alvo)	Fazer a análise semântica e o estudo piloto

Fonte: elaboração das autoras.

Como já mencionado, outras evidências de validade, como as de caráter empírico e estatístico, não são o foco deste capítulo. A seguir relatam-se os procedimentos metodológicos que foram utilizados para a criação da escala *Avaliação da Percepção de Competências Infocomunicacionais* (APCI).

RELATO

O objetivo principal do estudo foi criar e validar um instrumento para medir as competências infocomunicacionais dos estudantes da UFRGS, levando em conta a autoeficácia, uma característica psicossocial amplamente discutida na literatura da Psicologia, que aponta que a crença pessoal na própria capacidade influencia comportamentos, ações e decisões (Bandura, 1997; Bong, 2004; Mega; Ronconi; De Beni, 2014). Como explica Kurbanoglu (2003), a autoeficácia diz respeito à crença do indivíduo em sua própria capacidade de executar uma tarefa específica com êxito. Portanto, tem relação com a confiança pessoal.

SUMÁRIO

Em estudos da Ciência da Informação, essa crença demonstrou ter impacto direto na motivação para desenvolver competências infocomunicacionais (Clark, 2017). Estas, por sua vez, estão correlacionadas ao desempenho acadêmico (Bruce, 2004; Rosman; Anne-Kathrin; Krampen, 2015). Quando esta crença não está presente, as chances de subvalorização ou desinteresse pelo tema podem surgir.

O estudo teve natureza aplicada e pode ser considerado exploratório, pois serve de base para futuras pesquisas sobre o tema (Sampieri, 2014). Além disso, buscou solucionar um problema real, isto é, a falta de um instrumento para avaliar as percepções de autoeficácia em competências infocomunicacionais da comunidade usuária do Sistema de Bibliotecas da UFRGS (SBUFRGS). A essa comunidade é oferecido, desde 2017, um curso de promoção de competências infocomunicacionais chamado Super 8, voltado para a pesquisa e uso da informação científica. O curso utiliza um modelo baseado no Big 6, com o acréscimo de mais duas etapas ao processo de pesquisa: o *Comunicar* e o *Ser lido e avaliado*, por isso a denominação Super 8. O estudo é metodológico, pois focou na criação e validação de ferramentas de pesquisa (Moura; Gabriel Junior, 2019).

A abordagem de análise foi mista - quantitativa e qualitativa - utilizando o Coeficiente de Validade de Conteúdo (CVC) para a análise de juízes. As demais etapas - a análise semântica e o estudo piloto - foram qualitativas, coletando as opiniões dos participantes sobre o instrumento, a primeira via grupo focal e a segunda por meio da aplicação da escala em público-alvo.

Seguindo diretrizes psicométricas, a criação da escala APCI começou com uma extensa revisão da literatura, utilizando descritores em inglês e português e consultando diversas bases de dados acadêmicas da área da Ciência da Informação e da Psicologia. Modelos e padrões internacionais de competência, como ISP, Big 6, Framework e o próprio Super 8, foram fundamentais para definir a dimensionalidade do construto e os itens do instrumento.

SUMÁRIO

O quadro 2 demonstra um extrato da dimensionalidade do instrumento, suas definições constitutivas e as emoções envolvidas, considerando o processo de pesquisa de forma holística:

Quadro 2 – Dimensões da APCI

Dimensão	Definição constitutiva	Emoções
Reconhecer	Identificação do problema e planejamento da pesquisa, incluindo fontes ideais e formato preferido (digital, impresso ou outro)	Incerteza e ansiedade
Buscar	Localização inicial das fontes com critérios de confiabilidade e relevância	Ansiedade, declínio da confiança
Recuperar e Acessar	Seleção e armazenamento de fontes, utilizando ferramentas e filtros	Engajamento, interesse no tema
Usar	Interação com a informação para integrá-la ao conhecimento prévio, com estratégias como resumo, fichamento e mapas mentais para compreensão e apropriação	Diminuição da incerteza, mais confiança
Sintetizar	Reunião e organização da informação para formar interpretações baseadas em evidências e preparar o produto final, gerando uma visão personalizada	Alívio, satisfação ou frustração.
Comunicar	Definição de canais e estrutura visual para apresentação do trabalho de forma ética e adequada ao público.	-
Ser lido e avaliado	Utilização de estratégias para ampliar a visibilidade acadêmica, entendendo métricas e impacto científico.	-
Avaliar	Reflexão sobre o processo e o produto final, avaliando aprendizado, confiança, dificuldades e coesão do trabalho para aprimorar futuras pesquisas.	-

Fonte: elaboração própria com base na dissertação de Souza, 2022.

SUMÁRIO

A APCI foi avaliada por um grupo de cinco juízas, onde quatro eram bibliotecárias do SBUFRGS e uma psicometrista. As primeiras avaliaram conteúdo propriamente dito, utilizando os seguintes critérios sugeridos por Pasquali (1999):

- relevância teórica - os itens devem ser pertinentes ao construto;
- clareza - os itens devem ser compreensíveis para todos os estratos;
- pertinência - os itens são importantes para o instrumento.

A psicometrista avaliou, além dos critérios acima, a aplicação das regras de construção dos itens. A figura 1 demonstra um extrato do formulário de respostas das juízas:

Figura 1 – Formulário de respostas das juízas

Itens de Reconhecer	Relevância teórica					Clareza					Pertinência					Alterar redação para:	Transferir item para:	Sugestões e comentários
1 = pouquíssima, 2 = pouca 3 = média, 4 = muita 5 = muitíssima	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	Vai ser importante ter um bom rapport para os alunos responderem pensando em um mo		
Busco ideias e oportunidades investigando/buscando informações	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	☐	☐	☐	☐	☒			Não está muito claro... E contém duas informações no mesmo item.
Utilizo a literatura científica para identificar uma lacuna de conhecimento	☐	☐	☐	☒	☐	☐	☐	☐	☐	☒	☐	☐	☐	☐	☒			

Fonte: elaboração das autoras (2022).

As respostas foram analisadas e, de um universo de 301 itens, restaram 273 nesta etapa. A figura 2 demonstra um extrato da planilha com o índice CVC dos itens e a interpretação:

SUMÁRIO

Figura 2 – planilha com o índice CVC

Participante	Item 6			Item 7			Item 8		
	Relevância	Clareza	Pertinência	Relevância	Clareza	Pertinência	Relevância	Clareza	Pertinência
J1	4	4	4	5	5	5	5	5	5
J2	5	3	5	5	5	5	5	5	5
J3	5	2	5	5	5	5	5	5	5
J4	4	5	4	5	5	5	5	5	3
J5	3	5	4	5	5	5	5	5	5
No. da amostra	5								
CVC	0,840	0,760	0,880	1,000	1,000	1,000	1,000	1,000	0,920
Interpretação	Aceitável	Inaceitável	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável

Fonte: elaboração das autoras (2022).

Esta etapa foi concluída após mais uma rodada de avaliação com uma das juízas, que sugeriu um reagrupamento dos itens para fins de contextualização. Os itens foram organizados em categorias temáticas representando ações do processo de pesquisa. Em seguida, em uma atividade prática, os itens foram impressos, recortados e distribuídos em uma mesa, sendo agrupados novamente para reduzir itens similares ou que poderiam ser unificados, otimizando a sua quantidade. Restaram então 178 itens.

Na etapa de análise semântica, participaram estudantes de graduação e pós-graduação da UFRGS, inicialmente convidados em sessões do programa de capacitação do SBUFRGS - o Super 8, mas devido à baixa adesão, foram contatados alunos do Grupo de Pesquisa em Competências InfoComunicacionais (InfoCom). Em duas sessões de grupo focal, os participantes discutiram a compreensão dos itens, tamanho da APCI, linguagem e layout da escala, guiados pela pesquisadora e com discussões gravadas pela ferramenta de web conferência Mconf. Os dados foram transcritos, organizados em planilhas, relacionando os itens por dimensão, e analisados. Em suma, questões de vocabulário foram alteradas e mais algumas excluídas. Por exemplo, alguns alunos não sabiam o que significava a expressão “lacuna de conhecimento” presente em

SUMÁRIO

um dos itens. O grupo focal continha 8 alunos (7 de graduação e 1 de mestrado) de cursos diversos, sendo a maioria da Biblioteconomia. A composição do grupo é uma limitação do estudo, pois muitos dos participantes já tinham familiaridade com o campo estudado.

Por fim, o estudo piloto foi essencial para confirmar a aplicabilidade do instrumento, assegurando que ele fosse compreendido pelos respondentes sem a necessidade de intervenção do pesquisador. Os participantes responderam ao instrumento de forma online e deixaram seus comentários. Alunos da UFRGS e comunidade universitária externa participaram desta etapa, num total de 15 pessoas. Poucas sugestões foram feitas, mas, em resumo, algumas explicações foram acrescidas aos itens, como por exemplo o significado de *Pre-print*. Ao final, restaram 136 itens na escala. A seguir é tratada a questão da avaliação das competências infocomunicacionais.

AVALIAÇÃO

A avaliação é fundamental para garantir a efetividade dos programas de competências infocomunicacionais, permitindo que o público-alvo se aproprie da informação, gere conhecimento e o comunique de forma eficaz e ética. Existem três tipos de avaliação principais: a diagnóstica, feita antes do início do programa para detectar o nível de conhecimento dos alunos; a formativa, aplicada durante a formação para ajustar o conteúdo conforme as necessidades; e a somativa, feita no final para verificar a efetividade do aprendizado (Belluzzo; Santos; Almeida Júnior, 2014). Esses métodos permitem ajustar os programas e materiais instrucionais, de acordo com as necessidades específicas dos alunos.

A autoavaliação em forma de escala Likert foi a forma preferida para a criação da APCI. Esse método coloca o aluno no centro

SUMÁRIO

de sua aprendizagem. Autores como Arnone, Small e Reynolds (2010) defendem sua importância, argumentando que a autoavaliação pode aumentar a confiança e o engajamento dos alunos em relação ao aprendizado. Ela se adequa melhor à avaliação de atitudes, pois mobiliza conhecimentos e habilidades, algo que as avaliações tradicionais frequentemente ignoram. Aqui “mobilizar” significa que, quando o aluno se autoavalia, ele não está apenas lembrando o que aprendeu, mas também pensando em como usou, o que percebeu e sentiu sobre esse aprendizado. Por isso, autores como Walters *et al.* (2020) defendem que é necessário incluir medidas que avaliem o engajamento e a confiança dos alunos em relação às competências. Esse tipo de avaliação pode ser feito em qualquer momento: antes, durante ou depois de uma sessão ou programa de capacitação.

Embora útil para capturar as percepções dos alunos, a autoavaliação enfrenta críticas por sua dependência das percepções individuais, que podem estar distorcidas, pois algumas pessoas podem superestimar suas competências, enquanto outras podem fazer o oposto. Este fenômeno é chamado efeito Dunning-Kruger (Dunning, 2019; Kruger; Dunning, 2009). A avaliação ideal, no entanto, deve ser uma combinação de métodos que forneça um panorama abrangente das competências dos alunos e da eficácia dos programas, considerando tanto o aspecto técnico quanto o afetivo e atitudinal do aprendizado.

No Brasil, no entanto, ainda há carência de instrumentos psicometricamente adequados para avaliar competências infocomunicacionais (Pellegrini; Estácio; Vitorino, 2016; Santos, 2011; Santos; Casarin, 2014) reforçando a necessidade de seguir critérios científicos na construção desses instrumentos, uma vez que podem afetar decisões institucionais e individuais.

CONSIDERAÇÕES FINAIS

A avaliação dos programas de competências infocomunicacionais é um componente essencial para garantir a efetividade do aprendizado, promovendo ajustes contínuos nos conteúdos e métodos de ensino. Ao englobar diferentes tipos de avaliação — diagnóstica, formativa e somativa —, é possível atender às necessidades específicas dos alunos, bem como ajustar os programas para melhorar a apropriação da informação e o desenvolvimento de competências. Embora a autoavaliação enfrente críticas por possíveis distorções, ela é um método valioso para captar as percepções dos estudantes e aumentar seu engajamento.

A Psicometria, área dedicada à mensuração de atributos psicológicos, como atitudes e traços de personalidade, desempenha um papel crucial no desenvolvimento de instrumentos de avaliação confiáveis e válidos. Na Ciência da Informação, onde a avaliação das competências infocomunicacionais é fundamental, a aplicação da Psicometria pode elevar a qualidade dos testes utilizados para medir essas competências.

Ao utilizar princípios psicométricos, como a validade, é possível garantir que os instrumentos de avaliação realmente mensurem o que se propõem a medir e que o façam de forma consistente. Isso permite obter diagnósticos mais precisos tanto sobre o desempenho dos alunos quanto sobre a eficácia dos programas. Instrumentos bem construídos com base na Psicometria podem, por exemplo, detectar com maior exatidão as áreas em que os alunos enfrentam dificuldades, ajudando a moldar intervenções pedagógicas mais eficazes.

Isso é especialmente relevante na Ciência da Informação, que abrange um conjunto complexo de conhecimentos, habilidades e atitudes que vão além da simples aquisição de informação, incluindo sua interpretação, disseminação e uso ético. O avanço

SUMÁRIO

da Psicometria nesse campo pode fortalecer a base científica das avaliações, ajudando a Ciência da Informação a se consolidar ainda mais como uma área que alia teoria e prática de forma rigorosa e com métodos científicos reconhecidos.

A incorporação da Psicometria nos instrumentos de avaliação de competências infocomunicacionais, especialmente no contexto brasileiro, é ainda limitada, mas representa uma oportunidade significativa de crescimento, tanto na formação dos alunos quanto no aperfeiçoamento dos programas educacionais na área.

REFERÊNCIAS

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; AMERICAN PSYCHOLOGICAL ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. **Standards for educational and psychological testing**. Washington, D.C: American Educational Research Association, 2014. Disponível em: <https://pt.b-ok.lat/book/3489335/112112>. Acesso em: 4 fev. 2021.

ARNONE, Marilyn P; SMALL, Ruth V; REYNOLDS, Rebecca. Supporting inquiry by identifying gaps in student confidence: development of a measure of perceived competence. **School Libraries Worldwide**, Seattle, WA, v. 16, n. 1, p. 47-60, 2010. Disponível em: <https://surface.syr.edu/cgi/viewcontent.cgi?article=1094&context=istpub>. Acesso em: 29 jun. 2021.

BANDURA, Albert. **Self-efficacy: the exercise of control**. New York: W.H. Freeman, 1997. Disponível em: https://www.academia.edu/28274869/Albert_Bandura_Self_Efficacy_The_Exercise_of_Control_W_H_Freeman_and_Co_1997_pdf.

BELLUZZO, Regina Célia Baptista; SANTOS, Camila Araújo dos; ALMEIDA JÚNIOR, Oswaldo Francisco de. A competência em informação e sua avaliação sob a ótica da mediação da informação: reflexões e aproximações teóricas. **Informação & Informação**, Londrina, v. 19, n. 2, p. 60-77, 2014. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/19995>. Acesso em: 26 jun. 2021.

SUMÁRIO

BONG, Mimi. Academic motivation in self-efficacy, task value, achievement goal orientations, and attributional beliefs. **The Journal of Educational Research**, London, v. 97, n. 6, p. 287–298, 2004. Disponível em: <https://doi.org/10.3200/JOER.97.6.287-298>. Acesso em: 23 jul. 2020.

BORSA, Juliane Callegaro; SEIZE, Mariana de Miranda. Construção e adaptação de instrumentos psicológicos: dois caminhos possíveis. *In*: DAMÁSIO, Bruno Figueiredo; BORSA, Juliane Callegaro (org.). **Manual de desenvolvimento de instrumentos psicológicos**. São Paulo: Vetor, 2017. p. 15–37.

BRUCE, Christine. Information Literacy as a Catalyst for Educational Change: a background Paper. *In*: 3RD INTERNATIONAL LIFELONG LEARNING CONFERENCE, 2004, Yeppoon, Queensland. **Proceedings “Lifelong Learning: Whose responsibility and what is your contribution?”** Yeppoon, Queensland: [s. n.], 2004. p. 8–19. Disponível em: https://www.researchgate.net/publication/27466455_Information_Literacy_as_a_Catalyst_for_Educational_Change_A_Background_Paper. Acesso em: 26 nov. 2024.

CHANG, Yun-ke; ZHANG, Xue; MOKHTAR, Intan Azura; FOO, Schubert; MAJID, Shaheen; LUYT, Brendan; THENG, Yin-leng. Assessing students' information literacy skills in two secondary schools in Singapore. **Journal of Information Literacy**, United Kingdom, v. 6, n. 2, p. 19–34, 2012. Disponível em: <http://ojs.lboro.ac.uk/ojs/index.php/JIL/article/view/1694>. Acesso em: 9 dez. 2019.

CLARK, Melissa. Imposed-inquiry information-seeking self-efficacy and performance of college students: a review of the literature. **The Journal of Academic Librarianship**, Oxford, v. 43, n. 5, p. 417–422, 2017. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0099133316301689>. Acesso em: 11 jan. 2020.

CRAIG, Knight. The History of Psychometrics. *In*: CRIPPS, Barry (org.). **Psychometric Testing: Critical Perspectives**. Hoboken: John Wiley & Sons, Ltd, 2017. p. 1–14. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119183020.ch1>. Acesso em: 15 out. 2024.

DUNNING, David. The best option illusion in self and social assessment. **Self and Identity**, United States, v. 18, n. 4, p. 349–362, 2019. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/15298868.2018.1465460>. Acesso em: 8 nov. 2020.

GAMST, Glenn; MEYERS, Lawrence S.; BURKE, Holly McClain; GUARINO, A. J.; Scale development and validation. *In*: GUEST, Greg; NAMEY, Emily E. **Public Health Research Methods**. London: SAGE, 2017. p. 1–35. Disponível em: <http://methods.sagepub.com/book/public-health-research-methods>. Acesso em: 31 dez. 2020.

SUMÁRIO

GURUNG, Regan A. R. Best practices in scale use in SoTL. /n: JHANGIANI, Rajiv S.; TROISI, Jordan D.; FLECK, Bethany; LEGG, Angela M.; HUSSEY, Heather D. (ed.). **A compendium of scales for use in the scholarship of teaching and learning**. Washington, D.C: Society for the Teaching of Psychology, 2015. p. 11–19. Disponível em: https://www.academia.edu/25614591/A_COMPENDIUM_OF_SCALES_for_use_in_the_SCHOLARSHIP_OF_TEACHING_AND_LEARNING_EDITED_BY. Acesso em: 11 jul. 2021.

HAUCK FILHO, Nelson; ZANON, Cristian. Questões básicas sobre mensuração. /n: PSICOMETRIA. Porto Alegre: Artmed, 2015. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788582712368/cfi/6/8!/4/2/16@0:76.5>. Acesso em: 2 dez. 2020.

JONES, Lyle V.; THISSEN, David. A history and overview of Psychometrics. **Handbook of Statistics**, United States, v. 26, p. 27, 2007. Disponível em: [https://doi.org/10.1016/S0169-7161\(06\)26001-2](https://doi.org/10.1016/S0169-7161(06)26001-2). Acesso em: 19 nov. 2020.

KAREN B. GOLDFINGER. The Science of Psychological Testing: Fundamental Concepts. /n: GOLDFINGER, Karen B. **Psychological Testing in Everyday Life: History, Science, and Practice**. Thousand Oaks: SAGE Publications, Inc., 2019. p. 6–14. Disponível em: <https://methods.sagepub.com/book/psychological-testing-in-everyday-life>. Acesso em: 13 out. 2024.

KRUGER, Justin; DUNNING, David. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. **Psychology**, United States, v. 1, p. 30–46, 2009. Disponível em: /doiLanding?doi=10.1037%2F0022-3514.77.6.1121. Acesso em: 6 nov. 2020.

KURBANOGU, Serap. Self-efficacy: a concept closely linked to information literacy and lifelong learning. **Journal of Documentation**, Bingley, v. 59, n. 6, p. 635–646, 2003. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410310506295/full/html>. Acesso em: 11 jan. 2020.

LOPES, Carlos; PINTO, Maria. Desenho e construção do questionário de literacia da informação no ensino superior – IL-HUMASS: estudo qualitativo (Parte I). **Cadernos BAD**, Lisboa, n. 1/2, p. 123–139, 2013. Disponível em: <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1035>. Acesso em: 5 jan. 2020.

MAROCO, João; GARCIA-MARQUES, Teresa. Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas?. **Laboratório de Psicologia**, Lisboa, v. 4, n. 1, p. 65–90, 2006. Disponível em: <http://publicacoes.ispa.pt/index.php/lp/article/view/763>. Acesso em: 26 dez. 2020.

SUMÁRIO

MEGA, Carolina; RONCONI, Lucia; DE BENI, Rossana. What makes a good student? how emotions, self-regulated learning, and motivation contribute to academic achievement.

Journal of Educational Psychology, Washington, DC, v. 106, n. 1, p. 121-131, 2014.

Disponível em: /fulltext/2013-22762-001.html. Acesso em: 9 ago. 2020.

MOURA, Ana Maria Mielniczuk de; GABRIEL JUNIOR, Rene Faustino. **Metodologia da pesquisa científica**. Porto Alegre: [s. n.], 2019. Disponível em: https://moodle.ufrgs.br/pluginfile.php/3502824/mod_resource/content/0/Aula%2004%20-%20M%C3%A9todos%20de%20Pesquisa.pdf. Acesso em: 14 jul. 2021.

PACICO, Juliana Cerentini. Como é feito um teste? produção de itens. *In*: HUTZ, Claudio Simon; BANDEIRA, Denise Ruschel; TRENTINI, Clarissa Marceli. **Psicometria**. Porto Alegre: Artmed, 2015. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788582712368>. Acesso em: 2 dez. 2020.

PACICO, Juliana Cerentini; HUTZ, Cláudio; SCHNEIDER, Andreia M. de A.; BANDEIRA, Denise. Validade. *In*: PSICOMETRIA. Porto Alegre: Artmed, 2015. Disponível em: <https://app.minhabiblioteca.com.br/#/>. Acesso em: 11 nov. 2024.

PASQUALI, Luiz. Análise dos itens. *In*: PSICOMETRIA. Petrópolis: Vozes, 2017.

PASQUALI, Luiz (org.). **Instrumentação psicológica : fundamentos e práticas**. Porto Alegre: Artmed, 2010. Disponível em: <http://search.ebscohost.com/login.aspx?direct=true&db=catt07377a&AN=sabi.000746824&site=eds-live>. Acesso em: 4 dez. 2020.

PASQUALI, Luiz. Psicometria Psicometria Psychometrics. **Revista da Escola de Enfermagem da USP**, São Paulo, v. 43, n. (SPE), p. 992-999, 2009. Disponível em: http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0080-62342009000500002&lng=pt&nrm=iso&tlng=pt.

PASQUALI, Luiz. Testes referentes a construto: teoria e modelo de construção. *In*: INSTRUMENTOS PSICOLÓGICOS: MANUAL PRÁTICO DE ELABORAÇÃO. Brasília: LabPAM, 1999. p. 37-60.

PELLEGRINI, Eliane; ESTÁCIO, Letícia Silvana dos Santos; VITORINO, Elizete Vieira. Instrumentos de avaliação da competência em informação: um mapeamento em âmbito nacional e internacional. *In*: COMPETÊNCIA EM INFORMAÇÃO : POLÍTICAS PÚBLICAS, TEORIA E PRÁTICA. Salvador: EDUFBA, 2016. p. 155-182. Disponível em: <http://repositorio.ufba.br/ri/handle/ri/22598>. Acesso em: 28 jun. 2021.

SUMÁRIO

PETT, Marjorie; LACKEY, Nancy; SULLIVAN, John. Designing and testing the instrument. In: MAKING SENSE OF FACTOR ANALYSIS. Thousand Oaks: SAGE, 2003. p. 1–30. Disponível em: <http://methods.sagepub.com/book/making-sense-of-factor-analysis>. Acesso em: 28 dez. 2020.

ROSMAN, Tom; ANNEKATHRIN, Mayer; KRAMPEN, Günter. Intelligence, academic self-concept, and information literacy: the role of adequate perceptions of academic ability in the acquisition of knowledge about information searching. **Information Research: an International Electronic Journal**, Boras, v. 20, n. 1, p. 1–13, 2015. Disponível em: <http://informationr.net/ir/20-1/istic2/istic34.html>. Acesso em: 15 jul. 2021.

SAMPIERI, Roberto Hernández. **Metodología de la investigación**. 6. ed. Ciudad de México: McGraw-Hill, 2014. Disponível em: <http://observatorio.epacartagena.gov.co/wp-content/uploads/2017/08/metodologia-de-la-investigacion-sexta-edicion.compressed.pdf>. Acesso em: 27 jun. 2020.

SANTOS, Camila Araújo dos. **Análise de instrumentos de avaliação da competência informacional voltados para a educação superior**. 2011. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2011. Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/santos_ca_me_mar.pdf. Acesso em: 26 dez. 2020.

SANTOS, Camila Araújo dos; CASARIN, Helen Castro Silva. Habilidades informacionais abordadas em instrumentos de avaliação de CI. **Informação & Sociedade: Estudos**, João Pessoa, v. 24, n. 3, p. 135–144, 2014. Disponível em: <http://www.brapci.inf.br/index.php/res/v/92010>. Acesso em: 5 jan. 2020.

SARTES, Laisa Marcorela Andreoli; SOUZA-FORMIGONI, Maria Lucia Oliveira de. Avanços na Psicometria: da Teoria Clássica dos Testes à Teoria de Resposta ao Item. **Psicologia: Reflexão e Crítica**, Porto Alegre, v. 26, n. 2, p. 241–250, 2013. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79722013000200004&lng=pt&nrm=iso&tlng=en. Acesso em: 6 dez. 2020.

SELAU, Thais; SILVA, Mônia Aparecida da; BANDEIRA, Denise Ruschel. Construção e evidências de validade de conteúdo da Escala de Funcionamento Adaptativo para Deficiência Intelectual (EFA-DI). **Avaliação Psicológica**, Campinas, v. 19, n. 3, p. 333–341, 2020. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_abstract&pid=S1677-04712020000300012&lng=pt&nrm=iso&tlng=pt. Acesso em: 12 mar. 2021.

SUMÁRIO

SOUZA, Vanessa Inacio de. **Avaliação da percepção de competências infocomunicacionais (APCI): elaboração e busca de evidências de validade da escala**. 2022. 205 f. Dissertação (Mestrado em Ciência da Informação - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022). Disponível em: <http://www.bibliotecadigital.ufrgs.br/da.php?nrb=001150163&loc=2022&l=eee105a4d3c06935>. Acesso em: 4 nov. 2024.

URBINA, Susana. **Essentials of psychological testing**. 2. ed. Hoboken: Wiley, 2014. (Essentials of behavioral science). Disponível em: <https://pt.b-ok.lat/book/2361670/b99a3c>. Acesso em: 4 fev. 2021.

WALTERS, William; SHEEHAN, Sarah; HANDFIELD, Amy; LÓPEZ-FITZSIMMONS, Bernadette; MARKGREN, Susanne; PARADISE, Laurin. A multi-method information literacy assessment program: foundation and early results. **portal Libraries and the Academy**, Baltimore, v. 20, p. 101-135, 2020. Disponível em: Acesso em: 26 fev. 2020.

5

Ana Maria Mielniczuk de Moura

Fernanda Bochi

Janaina Lais Pacheco Lara Morandin

Maria Cláudia Cabrini Grácio

MÉTODOS DE PESQUISA APLICÁVEIS AOS ESTUDOS DE PATENTES NA CIÊNCIA DA INFORMAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.5

SUMÁRIO

COMO CITAR:

MOURA, Ana Maria Mieliczuk; BOCHI, Fernanda; PACHECO, Janaina Lais; MORANDIN, Lara; GRÁCIO, Maria Cláudia Cabrini. Métodos de Pesquisa Aplicáveis aos Estudos de Patentes na Ciência da Informação. /n: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações**. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

A Propriedade Intelectual (PI) é um ramo do Direito que tem como objetivo proteger as criações dos indivíduos e das organizações. Ela se divide em três categorias: Direito Autoral, que resguarda as criações artísticas, literárias e científicas; Propriedade Industrial, que abrange as patentes, marcas, desenhos industriais e indicações geográficas; e Proteção Sui Generis, destinada a casos específicos, como variedades vegetais e topografias de circuitos integrados. Neste capítulo, o foco está na Propriedade Industrial, com destaque para o tema das patentes.

As patentes representam um tema significativo de pesquisa na área da Ciência da Informação (CI), oferecendo dados para diversas análises e estudos, fornecendo uma compreensão profunda das dinâmicas de inovação tecnológica e ajudando a guiar decisões estratégicas em diversos níveis. Esses documentos são registros oficiais que protegem invenções, garantindo direitos exclusivos aos inventores por um período de 20 anos para patentes de invenção e 15 anos para modelos de utilidade (Moura *et al.*, 2020; Baltazar *et al.*, 2017). Esses documentos representam fontes primárias de informação técnica, oferecendo detalhes frequentemente ausentes em outras publicações científicas (Brito; Ozaki, 2020).

SUMÁRIO

A partir da perspectiva da epistemologia e da gestão da informação, as patentes são um desafio e uma oportunidade (Póvoa; Cabello, 2024). Nesse contexto, a análise de patentes desempenha um papel estratégico ao evitar a duplicação de esforços, permitindo o aproveitamento de soluções já existentes. Além disso, essas análises são ferramentas valiosas para monitorar a concorrência, possibilitando a identificação de empresas atuantes em setores tecnológicos específicos (Brito; Ozaki, 2020).

Realizada a partir dos dados que constam nos documentos, a análise de patentes pode revelar tendências tecnológicas, inovações emergentes e padrões de desenvolvimento industrial, proporcionando *insights* valiosos para pesquisadores, empresas e formuladores de políticas (Moura *et al.*, 2020; Baltazar *et al.*, 2017). A crescente quantidade de dados disponíveis em patentes pode exigir métodos avançados de recuperação e análise, bem como a adoção de big data, inteligência artificial e aprendizado de máquina. O uso dessas ferramentas não apenas ajuda a gerenciar a informação de maneira mais eficiente, mas também a descobrir padrões ocultos e prever inovações tecnológicas futuras (Póvoa; Cabello, 2024).

Ademais, a análise de patentes também tem implicações econômicas e estratégicas. Empresas podem usar essas informações para decidir onde investir em Pesquisa & Desenvolvimento (P&D), enquanto governos podem orientar políticas públicas de inovação. O estudo de redes de colaboração patentária pode revelar como o conhecimento é transferido entre diferentes atores, sejam eles empresas, universidades ou institutos de pesquisa. Além disso, a análise das citações às patentes pode indicar as tecnologias que estão em alta no mercado, podendo auxiliar pesquisadores a identificar tendências de pesquisa.

Ao conhecer o estado da técnica e os direitos já atribuídos, pesquisadores e empresas podem proteger suas inovações de forma mais eficaz, evitando litígios e promovendo um ambiente

SUMÁRIO

mais favorável à inovação. Assim, a Análise de Anterioridade é um procedimento essencial no campo da propriedade intelectual, especialmente no processo de patenteamento. Seu objetivo é verificar se uma invenção ou criação já foi divulgada anteriormente, garantindo que ela atenda aos critérios de novidade e atividade inventiva exigidos para a concessão de uma patente. Ao identificar previamente possíveis anterioridades, pode-se reduzir o risco de indeferimento do pedido, economizar tempo e recursos, além de realizar possíveis ajustes na invenção para garantir sua originalidade e relevância no mercado. Assim, a busca pela anterioridade torna-se uma etapa fundamental para confirmar a novidade da criação e viabilizar sua proteção legal (Linhares, 2017).

Compreender a relevância das patentes para o cenário econômico de um país, assim como para a indústria e as universidades, e quais métodos de pesquisa podem ser utilizadas para o estudo na área, reforça a pertinência deste capítulo, que tem como objetivo apresentar diversos métodos aplicáveis à pesquisa utilizando documentos de patente.

Assim, busca-se capacitar o leitor a extrair e analisar informações de patentes de forma eficaz, contribuindo para o avanço do conhecimento científico e tecnológico. São apresentados os conceitos de Propriedade Intelectual (PI), as características da patente e seus critérios de patenteabilidade, além dos elementos de um documento de patente que o qualificam como um instrumento relevante para a pesquisa, com destaque para os métodos aplicáveis aos estudos que tratam de patentes na CI.

SUMÁRIO

PROPRIEDADE INTELECTUAL (PI)

Antes de abordar a (PI), é pertinente refletir sobre o processo criativo humano que conduziu até o presente momento. A evolução do homem na pré-história abrange o desenvolvimento físico, social e cultural desde os primeiros hominídeos até o surgimento das primeiras civilizações. No período Paleolítico, a necessidade de sobrevivência levou ao desenvolvimento de ferramentas rudimentares de pedra lascada, ossos e madeira para caça e coleta. Com o tempo, a agricultura e a domesticação de animais no Neolítico exigiram instrumentos mais avançados. Além disso, povos antigos utilizavam métodos de contagem para registrar transações comerciais e eventos significativos, como as tábuas de contagem empregadas por egípcios e mesopotâmicos antes do desenvolvimento de sistemas formais de escrita. Os sumérios criaram a escrita cuneiforme, aproximadamente 3.200 a.C., visando atender às necessidades administrativas e econômicas de sua civilização. A escrita é uma das criações mais representativas que pavimentaram o caminho para o que se denomina Civilização (Paesani, 2012; Silveira, 2014).

Ao longo da história, a humanidade demonstra uma notável capacidade de desenvolver tecnologias que lhe permitem se adaptar e modificar o ambiente natural para atender às suas necessidades. Desde a confecção de ferramentas de pedra na pré-história até a criação de computadores e outros dispositivos avançados, o ser humano tem utilizado sua criatividade para superar desafios e aprimorar suas condições de vida. Essa trajetória evidencia uma relação intrínseca entre a evolução tecnológica e a capacidade humana de transformar o meio em que vive. Assim, conforme o ser humano aprimora suas habilidades tecnológicas, surge também uma busca por satisfação estética, visando atender não apenas às necessidades materiais, mas também ao enriquecimento espiritual. A arte e a tecnologia sempre estiveram interligadas, com avanços técnicos

SUMÁRIO

influenciando expressões artísticas ao longo da história. Em virtude de todo esse processo criativo, nasce a Propriedade Intelectual (Paesani, 2012; Silveira, 2014).

A PI é um conjunto de garantias concedidas, por lei, a pessoas físicas ou jurídicas sobre suas criações intelectuais. Esses direitos protegem obras e invenções advindas da criatividade humana, garantindo aos autores ou inventores o controle sobre o uso, a reprodução e a comercialização de suas criações por um período determinado. A PI é essencial para incentivar a inovação e o progresso tecnológico, cultural e econômico, oferecendo incentivos financeiros e reconhecimento aos criadores, ao mesmo tempo que contribui para o desenvolvimento da sociedade ao disponibilizar conhecimento e novas tecnologias. Consolidada a partir da Revolução Francesa, a Propriedade Intelectual aglutina três categorias, conforme a Figura 1.

Nessa perspectiva, observou-se que a PI precisou de um amparo legal, a partir de uma legislação nacional e das convenções internacionais. Em 1788, os Estados Unidos tornaram-se pioneiros em reconhecer e proteger a PI diretamente pela Constituição. Ou seja, naquele momento, a Propriedade Intelectual deixava de ser vista apenas como uma questão econômica, sendo incorporada ao ordenamento constitucional, com sua proteção sendo considerada um direito fundamental ou uma prioridade para o desenvolvimento da sociedade (Paesani, 2012).

SUMÁRIO

Figura 1 - Infográfico da Propriedade Intelectual



Fonte: elaborado pelas autoras.

No Brasil, a Constituição de 1988 reconhece a PI no artigo 5º, inciso XXVII, que afirma que “[...] os autores de obras científicas, artísticas ou literárias têm direito exclusivo de utilização, publicação ou reprodução de suas obras”.

No âmbito internacional, a PI está amparada pela:

- Convenção de Paris (1883): foi o primeiro acordo internacional voltado à proteção da propriedade industrial, abrangendo patentes, marcas, desenhos industriais e indicações geográficas. Estabeleceu princípios fundamentais, como o tratamento nacional e o direito de prioridade, permitindo que inventores e empresas solicitem proteção em outros países membros com base na data de depósito original.
- Convenção de Berna (1886): objetiva proteger os direitos dos autores sobre suas obras literárias e artísticas. Introduziu o conceito de proteção automática, garantindo que as obras sejam protegidas em todos os países signatários sem a necessidade de registro formal. Também estabeleceu o princípio do tratamento nacional, assegurando que autores estrangeiros recebam o mesmo tratamento que os nacionais.

SUMÁRIO

- Convenção de Roma (1961): estende a proteção aos direitos conexos ao direito autoral, abrangendo artistas intérpretes ou executantes, produtores de fonogramas e organizações de radiodifusão. Seu objetivo é assegurar que esses participantes do processo criativo tenham direitos sobre suas interpretações, gravações e transmissões.
- Conferência de Estocolmo (1967): Foi uma importante conferência que resultou na revisão e atualização de diversos tratados de propriedade intelectual, incluindo as Convenções de Paris e de Berna. Uma das principais realizações foi a criação da Organização Mundial da Propriedade Intelectual (OMPI), destinada a promover a proteção da PI globalmente.
- Organização Mundial da Propriedade Intelectual: é uma agência especializada das Nações Unidas dedicada a fomentar a proteção da propriedade intelectual no mundo. Administra diversos tratados internacionais e oferece uma plataforma para a cooperação entre países membros, visando harmonizar e aprimorar as leis e práticas de PI.
- Acordo TRIPS: Acordo sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio. Define padrões mínimos que os membros da Organização Mundial do Comércio (OMC) devem contemplar, entre eles:
 - direito autoral
 - marcas
 - indicações geográficas
 - desenhos industriais
 - patentes
 - circuitos integrados
 - informação confidencial

SUMÁRIO

A partir do elucidado, observa-se que a PI, enquanto objeto de estudo, integra de forma significativa à CI, uma área interdisciplinar que se dedica ao estudo da produção, organização, disseminação e uso da informação. Deste modo, este capítulo abordará a propriedade industrial, com ênfase nas patentes.

PATENTE DE INVENÇÃO E PATENTE DE MODELO DE UTILIDADE: CARACTERÍSTICAS

As patentes podem ser de dois tipos: patente de invenção e patente de modelo de utilidade, cujos níveis de invenção são diferenciados. Ambas são formas de proteção conferidas pelo Estado aos inventores para incentivar a inovação e o desenvolvimento tecnológico, porém, com diferenças importantes entre elas (Quadro 1).

Quadro 1 - Diferenças entre Patente de Invenção e Patente de Modelo de Utilidade

	Patente de Invenção	Patente de Modelo de Utilidade
Características	atividade inventiva; novidade; aplicação industrial	atividade inventiva; melhoria funcional no seu uso ou em sua fabricação
Prazo de proteção	20 anos a partir da data de depósito	15 anos a partir da data de depósito

Fonte: elaborado pelas autoras.

No caso da Patente de Invenção, a invenção não pode ter sido divulgada ou estar disponível ao público antes da data de depósito do seu pedido de patente, ou seja, não pode ter sido divulgada anteriormente de nenhuma forma, como publicações, exposições ou

SUMÁRIO

vendas. Deve também apresentar um grau de inventividade que não seja óbvio para um técnico especializado na área correspondente, assim, a invenção não pode ser uma solução evidente ou trivial. Como último aspecto, precisa ser passível de fabricação ou utilização em qualquer tipo de indústria, desde que dotada de repetibilidade (Instituto Nacional da Propriedade Industrial, 2021).

Já a Patente de Modelo de Utilidade precisa envolver um ato inventivo, porém suficiente no sentido de apresentar uma melhoria a uma invenção já existente, seja no seu uso ou na fabricação (Instituto Nacional da Propriedade Industrial, 2021).

Cabe ressaltar que existe ainda o Certificado de Adição de Invenção. Ele é um título acessório somente à patente de invenção, destinado a proteger aperfeiçoamentos ou desenvolvimentos introduzidos no objeto original. É importante mencionar que no sistema de propriedade industrial brasileiro, o Desenho Industrial é protegido por meio de um registro e não por uma patente. Isso ocorre porque o desenho industrial se refere aos aspectos ornamentais ou estéticos de um produto, como sua forma ou combinação de linhas e cores, que conferem uma aparência nova e original passível de aplicação industrial (Instituto Nacional da Propriedade Industrial, 2023).

O sistema de patentes estabelece um equilíbrio fundamental entre a proteção dos direitos do inventor e o interesse público no acesso à informação tecnológica. Ao conceder uma patente, o Estado oferece ao titular o direito exclusivo de explorar sua invenção por um período determinado — 20 anos para patentes de invenção e 15 anos para modelos de utilidade. Em contrapartida, o inventor é obrigado a divulgar detalhadamente sua criação no pedido de patente, permitindo que qualquer interessado acesse o conhecimento técnico envolvido. Essa divulgação possibilita que pesquisadores, empresas e outros inventores estudem a tecnologia patenteada, servindo de base para novas inovações e aprimoramentos. Dessa forma, os concorrentes podem desenvolver suas pesquisas, subsequentemente

SUMÁRIO

ao estágio mais avançado do conhecimento, promovendo o progresso tecnológico do país. Por esse sistema, não apenas se protege os direitos dos inventores, mas também assegura que a sociedade se beneficie do avanço contínuo da ciência e da tecnologia, incentivando novas descobertas e inovações futuras (Instituto Nacional da Propriedade Industrial, 2021).

DOCUMENTO DE PATENTE

No processo de patenteamento no Brasil, existem duas publicações principais que comunicam o status de uma invenção: o folheto de publicação e o folheto de concessão (Instituto Nacional da Propriedade Industrial, 2021). Conforme o Manual de Patentes (Instituto Nacional da Propriedade Industrial, 2021) existem dois tipos de folhetos, um denominado de folheto de publicação e outro de folheto de concessão. Eles são formados por uma folha de rosto que contém os dados de identificação do pedido; a saber: nome do depositante, inventor, data de depósito, número da patente, resumo e demais informações contidas no documento.

- A. **Folheto de Publicação:** Após a submissão de um pedido de patente e a aprovação no exame formal preliminar, o Instituto Nacional da Propriedade Industrial (INPI) divulga o folheto de publicação. Este documento torna pública a invenção, permitindo que a sociedade conheça os detalhes técnicos do pedido depositado no Brasil. A publicação inicial geralmente ocorre 18 meses após a data de depósito ou da prioridade mais antiga reivindicada.
- B. **Folheto de Concessão:** Após a análise técnica detalhada e o deferimento do pedido, o INPI emite a carta-patente, oficializando a concessão dos direitos ao inventor.

SUMÁRIO

Neste momento, é publicado o folheto de concessão, que confirma que o pedido foi aprovado e que a patente está em vigor. Este documento inclui informações como o número da patente, data de concessão e detalhes sobre o titular.

Essas publicações são identificadas por códigos específicos, denominados *kind codes*. Por exemplo, para uma patente de invenção, o folheto de publicação é identificado pelo código "A2", ao passo que o folheto de concessão utiliza o código "B1". Esses códigos ajudam a distinguir o estágio em que o pedido de patente se encontra. Assim, o folheto de publicação divulga ao público os detalhes de uma invenção após a submissão do pedido, enquanto o folheto de concessão confirma a oficialização da patente após a conclusão bem-sucedida do processo de exame.

A Figura 2 ilustra a folha de rosto de um documento de patente, com o conjunto de informações nelas contidas. Usualmente, as análises de patentes utilizam dados extraídos dessas folhas de rosto para mapear o desenvolvimento tecnológico, seja por área, por país, instituição, inventores ou depositantes, entre outros. Nela, observam-se a data de publicação, registrada no código 43, país de depósito, informado no código 19, o número da patente (código 10), o nome do depositante (registrado em 71) e do inventor (em 72), a descrição resumida do item a ser patenteado (em 57), entre outras informações. Destaca-se que, nesse exemplo, depositante e inventor, respectivamente em 71 e 72, coincidem, mas eles podem diferir (figura 2).

SUMÁRIO

Figura 2 - Folha de rosto de uma patente depositada nos Estados Unidos (19) em 17 de agosto de 2017 (43: Pub. Date) relativa a um dispositivo para juntar dois instrumentos de escrita (54)


US 20170232781A1

(19) **United States**

(12) **Patent Application Publication**

(43) **Pub. No.: US 2017/0232781 A1**

Nguyen

(43) **Pub. Date: Aug. 17, 2017**

(54) **DEVICE FOR JOINING TWO WRITING INSTRUMENTS**

(71) Applicant: **Andrew Nguyen**, Huntington Beach, CA (US)

(72) Inventor: **Andrew Nguyen**, Huntington Beach, CA (US)

(21) Appl. No.: **15/041,680**

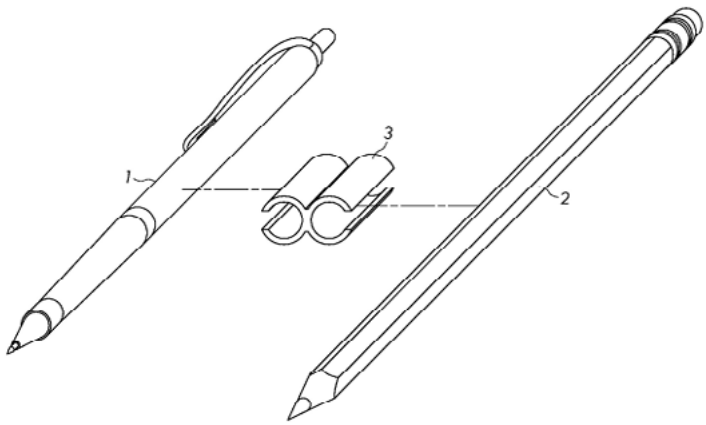
(22) Filed: **Feb. 11, 2016**

Publication Classification

(51) Int. Cl. **B43K 23/06** (2006.01)
F16B 2/22 (2006.01)

(52) U.S. Cl. **B43K 23/06** (2013.01); **F16B 2/22** (2013.01)

(57) **ABSTRACT**
This invention provides a device for removably connecting two different styles of writing instruments, such as pens with different colored inks or different size nibs, pencils, markers, or crayons.



Fonte: Nguyen (2017).

O Quadro 2 apresenta o tipo de informação a ser fornecida para cada código INID (Internationally agreed Numbers for the Identification of (bibliographic) Data) presente em uma patente, em termos de classe e tipo de informação.

Quadro 2 - Informações encontradas em documentos de patentes e seus códigos INID

Natureza	Número código INID	Tipo de informação	Descrição
Documento de patente	(10)	<i>Número de identificação do documento de patente</i>	É um identificador único atribuído a cada pedido de patente. É fundamental para a recuperação precisa de documentos de patente, garantindo que cada invenção ou pedido possa ser identificado de forma inequívoca.
	(12)	<i>Tipo do documento</i>	Indica, de forma detalhada, a natureza do documento, como uma publicação de pedido de patente ou uma patente concedida. Essa especificação facilita a identificação rápida do conteúdo e do estágio do processo de patenteamento.
Depósito de pedido de patente	(21)	<i>Número de identificação do depósito</i>	Código único designado ao pedido de patente no momento de seu depósito. Permite recuperação unívoca do documento.
	(22)	<i>Data do depósito</i>	Informação crucial para determinar o início do período de proteção da patente, que, em caso de concessão, geralmente é de 20 anos a partir dessa data para patentes de invenção.
Prioridades	(31)	<i>Número de identificação da prioridade</i>	Código único relativo aos documentos de prioridade, utilizados para reivindicar período de proteção com data de início anterior à data de depósito do pedido de patente.
	(32)	<i>Data do depósito da prioridade</i>	Informação utilizada no lugar da data de depósito do pedido de patente para cálculo de período de vigência de proteção (20 anos após essa data).
Publicação	(43)	<i>Data de publicação por impressão (ou meios similares) de pedidos de patentes (ainda) não concedidas</i>	Data relativa à publicação de pedidos de patente em análise, rejeitados ou abandonados. Define a data a partir da qual o documento deve ser considerado estado da técnica, incluindo a divulgação impressa ou digital.
	(45)	<i>Data de publicação por impressão (ou meios similares) de patentes.</i>	Data de publicação da concessão de uma patente, incluindo a divulgação impressa ou digital.

SUMÁRIO

Natureza	Número código INID	Tipo de informação	Descrição
Informações técnica	(51)	<i>Classificação internacional de patentes</i>	Código relativo às classificações internacionais IPC e CPC. Permite verificar o campo de aplicação da invenção. Ótimo parâmetro de busca para identificação de invenções semelhantes.
	(52)	<i>Classificação nacional</i>	Classificação do país onde a patente se encontra depositada. Utilizado principalmente por patentes estadunidenses.
	(54)	<i>Título da invenção</i>	Título designado à invenção.
	(56)	<i>Lista de documentos do estado da técnica</i>	Número de identificação de documentos de patente que constituem o estado da técnica citados.
	(57)	<i>Resumo</i>	Resumo da invenção, geralmente descrevendo de forma breve do que se trata a invenção, sua aplicação e características diferenciais. Útil na avaliação preliminar da invenção.
Pessoas físicas e jurídicas	(71)	<i>Nome do depositante</i>	Pessoa responsável pelo depósito do pedido de patente.
	(72)	<i>Nome do inventor</i>	Pessoa responsável por desenvolver a invenção. Geralmente uma pessoa física.
	(73)	<i>Nome do Titular</i>	é a pessoa ou entidade que detém os direitos legais sobre a invenção, podendo ser uma pessoa física ou jurídica, como empresas ou universidades. Em alguns casos, o titular pode ser distinto do depositante, que é a pessoa que formalmente apresenta o pedido de patente ao escritório
	(75)	<i>Nome do inventor/ depositante</i>	Usado em casos cujo inventor é o próprio depositante do pedido de patente.
	(76)	<i>Nome do inventor/ titular</i>	Usados em casos cujo inventor é também o aplicante e titular da patente.

Fonte: Elaborado pelas autoras (adaptado da WIPO, 2024).

SUMÁRIO

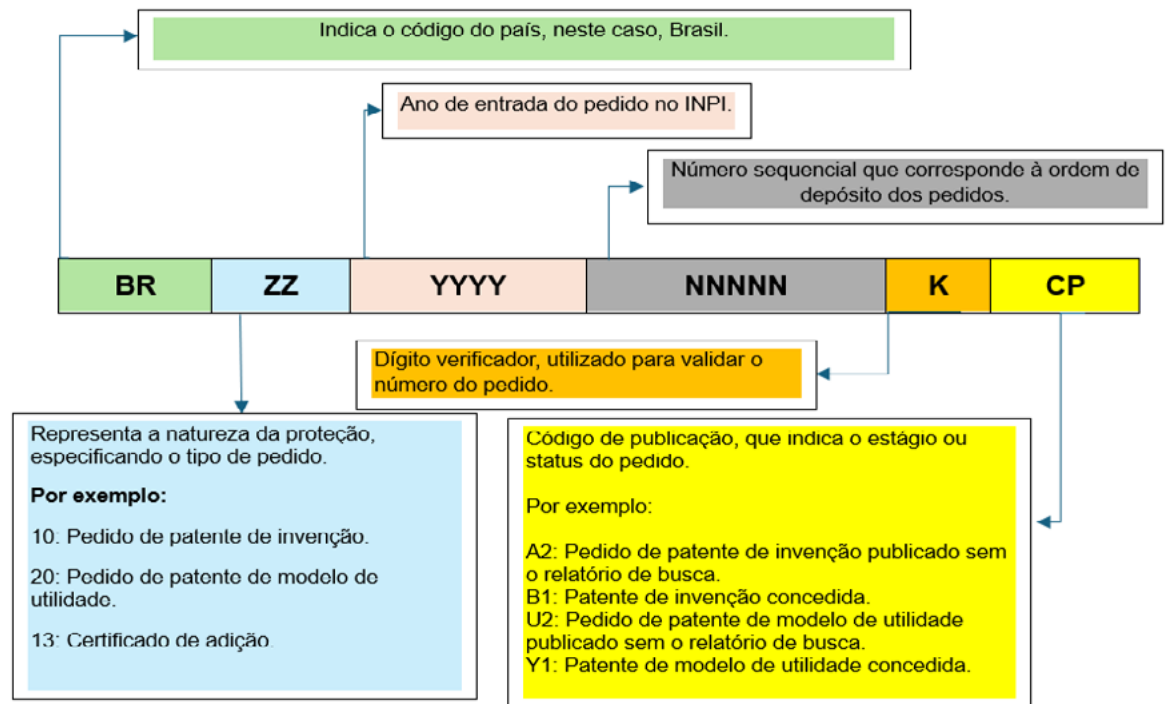
É importante compreender que o número de uma patente é composto por uma sequência de caracteres que fornece informações específicas sobre o pedido ou concessão (Espacenet, 2017). Há três elementos principais:

- A. Código do País:** Representado por duas letras que indicam o país onde o pedido foi efetuado.
- B. Ano de Depósito:** Consiste em quatro dígitos que indicam o ano em que o pedido foi realizado.
- C. Número Sequencial:** É um número único atribuído ao pedido dentro do ano correspondente.

A figura 3 ilustra a decodificação do número da patente em termos dos elementos descritos. Observa-se que o número da patente é formado por 16 dígitos; em que: a primeira informação é relativa ao código do país (na ilustração, BR relativa ao Brasil), seguida pela natureza da proteção, formada por dois dígitos (ZZ), acompanhado na sequência pelo ano de entrada do pedido (4 dígitos - YYYY), depois a ordem de depósito (5 dígitos - NNNNN) e o dígito verificador (1 dígito K). A última informação do número da patente, formado por 2 dígitos, diz respeito ao status do pedido (A2, B1, U2, Y1).

SUMÁRIO

Figura 3 - Decodificação do número de patente



Fonte: elaborado pelas autoras.

BASE DE DADOS DE PATENTES

De acordo com Bochi, Moura e Gabriel Junior (2020), ao conduzir pesquisas utilizando documentos de patentes, é fundamental compreender as diversas bases de dados disponíveis e as particularidades de cada uma. Por isso, os autores apresentaram, de forma breve, algumas delas e suas características, como é possível ver na figura 4.

Figura 4 - Principais bases de dados e suas características

Base de patente	Acesso	Cobertura	Em português	Gera análises	Responsável
Derwent Innovation Indexa (DII)	\$	Internacional		✓	Clarivate
Espacenet	✓	Internacional			European Patent Office
Google Patents	✓	Internacional	✓		Google
Base de dados do INPI	✓	Brasil	✓		INPI (Brasil)
LatiPat	✓	América Latina e Espanha			European Patent Office
Orbit	\$	Internacional		✓	Questel
Patentscope	✓	Internacional			Wipo
Lens.org	✓	Internacional		✓	Queensland University of Technology & Cambia
Free Patents Online	✓	EUA			SumoBrain
USPTO	✓	EUA			EUA

Fonte: Bochi, Moura e Gabriel Júnior (2020, p. 279).

SUMÁRIO

PRINCIPAIS MÉTODOS DE PESQUISA APLICÁVEIS AOS DOCUMENTOS DE PATENTE

Nesta seção, são apresentados os principais métodos de pesquisa aplicados nos estudos que envolvem os documentos de patentes, de forma a dar destaque a estes métodos, suas características e como podem ser aplicados, incluindo as ferramentas adequadas para esta aplicação. Os estudos quantitativos e qualitativos estão descritos separadamente, entretanto, ressalta-se que ambos dialogam nas pesquisas com patentes, de modo que cada abordagem contribui para o enriquecimento e a compreensão proporcionados pela outra.

MÉTODOS QUANTITATIVOS

Os métodos quantitativos caracterizam-se pelo uso de indicadores, métricas e análises estatísticas para interpretar fenômenos de forma objetiva e mensurável. No campo dos estudos de patentes, destacam-se:

ESTUDOS MÉTRICOS DA INFORMAÇÃO

Os Estudos Métricos da Informação (EMI), muitas vezes utilizados como sinônimo de Bibliometria, é um campo que utiliza métodos quantitativos para estudar a informação em diferentes contextos, incorporando subcampos como Bibliometria, Cientometria e Patentometria. É um guarda-chuva teórico-metodológico aplicável em diferentes cenários, desde a análise de publicações científicas até

SUMÁRIO

o impacto de tecnologias. Isso contribuiu para ampliar e consolidar o campo de estudos métricos, tornando-o uma área interdisciplinar que dialoga com a ciência da informação, a sociologia, a economia e a ciência da computação entre outras.

Grácio (2020) considera que os EMI são um campo que envolve pesquisa a respeito de análise e avaliação da informação, de caráter formal e informal, sendo sustentado por várias outras disciplinas como Ciência da Informação, Sociologia, Matemática, Estatística e Computação. Essas disciplinas, combinadas, utilizam procedimentos quantitativos como método para análise, ou seja, empregam métricas, estatísticas e modelos computacionais para interpretar fenômenos de forma objetiva e sistemática.

Neste sentido, uma variedade de indicadores são aplicados, possibilitando mensurar, por exemplo, produtividade e relações, tanto científicas quanto tecnológicas. Para Mugnaini, Jannuzzi e Quoniam (2004) os indicadores em C&T, podem se relacionar a insumos, como os investimentos, assim como à produtividade, por meio das publicações científicas e patentes registradas. Também podem medir o impacto científico e a capacidade de inovação das organizações.

No que tange às patentes, o acesso público a bases eletrônicas facilitou a análise de um grande conjunto de dados, o que permitiu um significativo avanço na compreensão dos indicadores (Hinze; Schmoch, 2004). Desta forma, as patentes se tornaram o indicador de produção de pesquisa aplicada mais utilizado (Hullmann; Meyer, 2003).

Dentre os métodos utilizados nos estudos métricos da informação, o que mais se aplica aos estudos de patentes é a Patentometria, discorrida a seguir.

SUMÁRIO

PATENTOMETRIA: ANÁLISE DOS DADOS DAS PATENTES E A PROSPECÇÃO TECNOLÓGICA

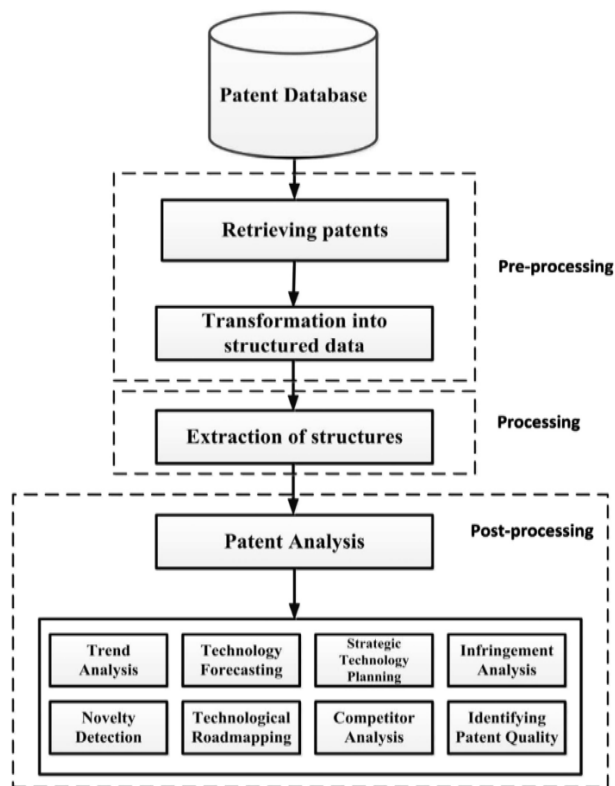
Os documentos de patentes são fontes valiosas para avaliar o desenvolvimento econômico e tecnológico de um país. Eles oferecem vantagens significativas em relação a outras fontes de informação, pois contêm dados bibliográficos e técnicos que auxiliam na resolução de problemas de pesquisa e desenvolvimento. Além disso, esses documentos refletem o estado da arte, ou seja, o nível mais avançado de desenvolvimento em um determinado campo técnico, sendo utilizados para determinar a originalidade e a inovação de uma invenção. Conforme destaca Guzmán Sánchez (1999), a análise de patentes, conhecida como Patentometria, é uma ferramenta eficaz para identificar oportunidades tecnológicas e medir a competitividade em C&T.

Park *et al.* (2013) ressaltam que as análises estatísticas realizadas a partir de dados patentários geram inteligência de patentes, isto é, transformação do conteúdo contido nos documentos em informações técnicas, comerciais e jurídicas. A análise aprofundada dos dados de patentes fornece suporte à tomada de decisão sobre o planejamento nacional de pesquisa, inovação e desenvolvimento, além de auxiliar na seleção de projetos tecnológicos (Aaldering; Song, 2019). A prospecção tecnológica em dados patentários deve ser realizada de forma rotineira a fim de influenciar processos de tomadas de decisão, uma vez que facilitam a apropriação da propriedade intelectual, melhoram a gestão da inovação, aumentam o senso crítico de Instituições Científicas, Tecnológicas e de Inovação (ICT), ampliam a visão dos gargalos tecnológicos e das oportunidades relacionadas a eles (Quintella *et al.*, 2011).

SUMÁRIO

Assim, é importante compreender esta abordagem como um processo de estágios múltiplos em que métodos e ferramentas são empregadas para obter resultados importantes com o objetivo proposto. Moehrle *et al.* (2010) dividem o processo em três estágios: i) pré-processamento (coleta e tratamento dos dados); ii) análise dos dados (aplicação de métodos e ferramentas para insights); iii) descoberta do conhecimento (análise dos resultados de forma visual e mensurável). Abbas, Zhang e Samee (2014) propuseram um processo similar (figura 5), em que as etapas se desdobram em pré-processamento, processamento e pós-processamento.

Figura 5 - Fluxo da pesquisa patentométrica



Fonte: Abbas; Zhang; Samee (2014).

SUMÁRIO

A etapa de pré-processamento é realizada pela definição das palavras-chave que definem o assunto a ser pesquisado, seguida pela escolha de uma base de patentes que atenda aos múltiplos aspectos de uma pesquisa, como a cobertura geográfica, por exemplo. A coleta dos dados implica na escolha dos campos que pretende-se analisar, o que inclui verificar a disponibilidade dos mesmos, assim como as formas de exportação de dados permitida pela base, automática ou não. Em seguida os dados precisam ser tratados, o que implica na padronização e limpeza dos mesmos.

A análise dos dados ocorre com a aplicação dos métodos escolhidos. Em se tratando de Patentometria, indicadores métricos são utilizados e necessitam de uma interpretação qualitativa que justifique ou explique os resultados encontrados.

A etapa de análise dos dados é ampla, abarcando a caracterização das patentes, como, por exemplo, evolução temporal, identificação de áreas temáticas e países protagonistas. A colaboração também pode ser observada, entre depositantes e entre inventores.

PROSPECÇÃO TECNOLÓGICA

Prospecção tecnológica é um método sistemático para identificar e mapear o desenvolvimento científico e tecnológico no futuro. O objetivo é antecipar as mudanças tecnológicas e identificar novas tecnologias emergentes. Também tem como objetivos antecipar as direções e velocidades das mudanças tecnológicas, detectar oportunidades e necessidades para a pesquisa e desenvolvimento e melhorar as decisões.

A análise bibliométrica de publicações científicas e patentes é uma abordagem eficaz para a prospecção tecnológica, permitindo mapear o desenvolvimento científico e identificar tendências emergentes.

SUMÁRIO

No estudo de Lee, Lee e Liaw (2012), os autores utilizaram a análise de publicações científicas para mapear o desenvolvimento tecnológico no setor agrícola de Taiwan, demonstrando como essa abordagem pode orientar estratégias de pesquisa e desenvolvimento. Embora focado em publicações científicas, os métodos empregados podem ser adaptados para a análise de patentes, oferecendo *insights* valiosos para identificar oportunidades emergentes e orientar políticas tecnológicas. Por outro lado, o trabalho de Procaci *et al.* (2016), exemplifica a aplicação direta da análise de patentes para mapear inovações tecnológicas no contexto educacional brasileiro, demonstrando como a prospecção baseada em patentes pode revelar tendências e direcionar estratégias de pesquisa e desenvolvimento. Ambos os estudos evidenciam a importância da análise de documentos científicos e patentários na identificação de tendências tecnológicas e na formulação de políticas de inovação.

Os métodos de prospecção tecnológica baseados em dados de patentes oferecem uma perspectiva quantitativa para identificar tendências e oportunidades emergentes. Entre as técnicas utilizadas destacam-se:

- A. **Análise de citações:** examina as referências entre patentes para mapear a evolução tecnológica e identificar inovações-chave.
- B. **Modelagem de tópicos:** utiliza algoritmos para detectar padrões e temas recorrentes em grandes volumes de textos de patentes.
- C. **Análise estatística:** aplica métodos estatísticos para quantificar tendências e correlacionar dados patentários com indicadores de mercado.
- D. **Análise de conteúdo:** envolve a interpretação detalhada das informações textuais nas patentes para compreender o contexto e as aplicações das inovações.

SUMÁRIO

CITAÇÃO DE PATENTES

A citação é um elemento central no processo de construção do conhecimento científico, que se manifesta sempre que um pesquisador X menciona, em sua publicação, o trabalho de outro pesquisador Y. A citação reflete o reconhecimento que o autor Y recebe do autor X pelo uso de seu conteúdo. Em contrapartida, a referência representa o reconhecimento dado pelo pesquisador X ao citar o conhecimento previamente publicado por Y (Egghe; Rousseau, 1990).

Nesse sentido, as referências estabelecem a ligação de uma pesquisa com seus fundamentos teóricos e metodológicos, ao passo que as citações evidenciam a relevância e o impacto de um trabalho na produção científica subsequente (Grácio, 2020).

No âmbito dos estudos de patentes, a análise de citação permite mapear o impacto das publicações científicas e das patentes no desenvolvimento de novas patentes e vice-versa, isto é, o impacto de uma patente em novas publicações científicas. Por meio dela, é possível inferir o impacto de um estudo ou patente, analisando a frequência com que são citados por estudos ou patentes subsequentes. Um número elevado de citações, portanto, indica que o estudo ou patente exerceu influência significativa em pesquisas ou patentes posteriores, evidenciando sua relevância para as comunidades científica e tecnológica.

A análise de citação tem como princípio básico a contagem das referências mencionadas ao final do documento, na lista de referências. Nas pesquisas desenvolvidas com a produção científica é possível identificar frentes de pesquisa, os principais autores, países, instituições, apontar conceitos, objetos e métodos, facilitando assim a compreensão das comunidades científicas (Vanz; Caregnato, 2003). Ajustada ao escopo da análise de patentes, a análise de citação permite identificar as tecnologias (patentes)

SUMÁRIO

mais citadas, os inventores que compõem a frente tecnológica de um campo, as empresas e suas relações com outras instituições.

Chang, Wu e Leu (2010) destacam que as citações em documentos de patentes, tanto a outras patentes quanto à literatura científica, são fundamentais para a análise de citações. Os autores argumentam que patentes frequentemente citadas indicam sua relevância e estabelecem a base para avanços tecnológicos e inovações. Além disso, as citações à literatura não patente evidenciam uma forte conexão entre tecnologia e ciência, demonstrando a fundamentação teórica presente nesses documentos.

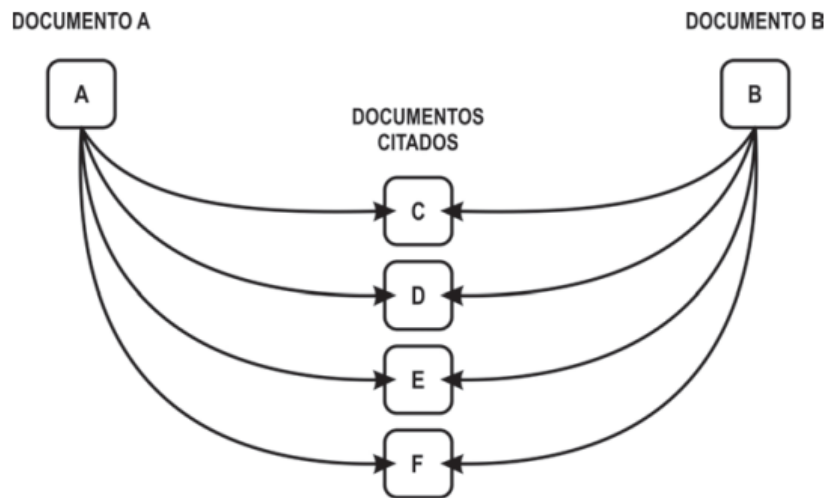
Assim, as citações em documentos de patentes possuem características distintas das encontradas na literatura científica. Tanto inventores quanto examinadores podem inserir referências em patentes, porém seus objetivos diferem. Marco (2007) observa que as citações em patentes indicam sua qualidade, medem o fluxo de conhecimento e permitem investigar o comportamento estratégico das empresas. Ferreira (2012) acrescenta que as patentes incluem referências a outras patentes e à literatura não patenteada, como livros, teses, artigos científicos, manuais e trabalhos apresentados em eventos, com o intuito de comprovar a originalidade da invenção. As citações presentes no relatório descritivo podem ser inseridas tanto pelo inventor quanto pelo examinador do pedido de patente. No entanto, há evidências de que as citações feitas pelo examinador durante a busca de anterioridade visam apresentar precedentes que possam refutar a invenção por falta de novidade.

As análises de citação de patentes podem ser realizadas observando as citações individualmente, com o intuito de relacionar o documento a outras patentes ou artigos científicos, conhecendo sua identidade. Também podem ser feitas análises relacionais, buscando estabelecer um relacionamento entre duas patentes. Assim, pode ser utilizado o Acoplamento Bibliográfico (figura 6), que permite identificar quando duas patentes compartilham pelo menos

SUMÁRIO

uma citação e a Análise de Cocitações, que identifica patentes que foram citadas juntas, pelo menos uma vez, por dois documentos.

Figura 6 - Documentos A e B acoplados bibliograficamente, em função dos quatro documentos C, D, E, F citados, em comum



Os documentos A e B estão acoplados porque ambos citam os documentos C, D, E e F

Fonte: Mattos e Dias (2010, adaptada de Garfield, 2001).

As análises de Acoplamento Bibliográfico e Cocitações podem ser feitas utilizando as linguagens de programação Python e R. Para a criação de redes de cocitação, o software VosViewer é uma opção amplamente utilizada, reconhecida por sua interface intuitiva e capacidade de gerar visualizações claras de redes bibliométricas. Além disso, o CiteSpace é outra ferramenta que permite a construção de redes de cocitação, oferecendo uma variedade de parâmetros para personalização das visualizações, embora sua interface possa ser considerada menos amigável. Recentemente, foi desenvolvida a ferramenta The Coupler, construída em R, que se destaca por ser gratuita e focada em análises relacionais de citação, incluindo acoplamento bibliográfico e cocitação.

SUMÁRIO

CO-OCORRÊNCIA DE PALAVRAS

A análise de coocorrência de palavras permite identificar os atores e explicar a dinâmica global da ciência (Callon; Courtial; Laville, 1991). Esta técnica considera a dinâmica da ciência como o resultado das estratégias dos atores. As mudanças nos conteúdos de uma área temática são o efeito combinado de muitas estratégias individuais. Robredo e Cunha (1998) afirmam, ainda, que a análise de co-ocorrência de palavras é uma técnica que avalia a frequência com que pares de termos aparecem juntos em um conjunto de documentos. Ao calcular índices estatísticos que representam a força de associação entre esses pares, é possível mapear o estado atual de uma área do conhecimento. Essa abordagem permite identificar clusters de palavras-chave e analisar a intensidade das conexões entre termos, facilitando a compreensão das estruturas conceituais e temáticas de um domínio específico. Além disso, auxilia na construção de léxicos especializados e no desenvolvimento de ferramentas para otimizar processos de indexação automática e recuperação da informação.

Um dos problemas que se encontra quando se faz uma análise de copalavras é a carência de normalização no uso das palavras-chave. As palavras-chave variam de uma base de dados a outra para os mesmos assuntos; também as usadas de um periódico a outro variam substancialmente para os mesmos assuntos, igualmente o uso das palavras variam de um autor a outro para os mesmos assuntos. Como não existe normalização das palavras-chave, existe também um total desconhecimento das normas de elaboração de palavras-chave e essas são uma exibição de variações na forma de uso: variam de plural para singular, profusão de sinônimos ou quase sinônimos, variações de campo a campo, e até desconhecimento do que seria propriamente uma análise bibliométrica. Portanto, antes de começar qualquer análise de copalavras seria recomendável normalizar o uso das palavras-chave (Urbizagástegui-Alvarado, 2022).

SUMÁRIO

Uma forma de aplicar a técnica de co-ocorrência de palavras é o uso da linguagem de programação R. Baseado nos valores da densidade e da centralidade, o pacote MPA do R constroi um diagrama estratégico que permite visualizar a maturidade e coesão dos assuntos pesquisados, assim como foi apresentado no estudo de Urbizagástegui-Alvarado (2022). No contexto das patentes, essa abordagem facilita a análise de grandes volumes de dados textuais presentes em documentos de patentes, oferecendo insights valiosos sobre as tendências e inter-relações em áreas tecnológicas específicas.

Assim, a análise de co-ocorrência de palavras em patentes é uma ferramenta poderosa para mapear e entender a evolução de campos científicos e tecnológicos. Ao superar desafios relacionados à padronização de termos e utilizando ferramentas analíticas adequadas, é possível extrair informações estratégicas que apoiam a tomada de decisões em pesquisa, desenvolvimento e inovação.

Além da ferramenta R, a análise de co-ocorrência de palavras pode ser feita por meio da linguagem de programação Python, que permite análises avançadas. Também a partir do Python, podem ser geradas redes usando o software Gephi. O NVivo, voltado às análises qualitativas também permite a criação de nuvens de palavras.

ANÁLISE DE REDES SOCIAIS

A Análise de Redes Sociais (ARS) é o conjunto de estudos destinado a investigar um fenômeno social por meio da análise das relações entre as entidades sociais, descrevendo os padrões e suas implicações (Wasserman; Faust, 1994). Permite visualizar as relações existentes em um conjunto de entidades sociais a fim de caracterizar e descrever a estrutura vigente (Grácio; Oliveira, 2020).

SUMÁRIO

Nesses estudos, focam-se as relações (também chamadas de ligações ou links) entre as unidades (entidades) de um grupo para compreender a dinâmica entre eles, com os atributos das entidades, como elementos auxiliares para o amplo entendimento das características das relações estabelecidas e do fenômeno estudado (Otte; Rousseau, 2002). Assim, o relacionamento entre um par de unidades é uma propriedade do par e não uma característica própria das entidades. Desse modo, a unidade básica de análise é medida em função da interação (relação) entre um par de entidades (Grácio; Oliveira, 2020).

Nos estudos bibliométricos, a ARS constitui valiosa contribuição metodológica para as pesquisas destinadas a mapear as coautorias científicas, redes de citação, de cocitação e acoplamento bibliográfico, assim como as co-ocorrências de palavras, entre outras relações presentes na atividade científica, possibilitando a visualização e compreensão das interações presentes na estrutura de uma comunidade ou campo científico (Otte; Rosseau, 2002).

Também nos estudos de patentes, a ARS vem se consolidando como método de análise das relações presentes entre seus elementos componentes, para redes de coinventores, de copaisés, de cocitação, assim como mapas de coclassificação de patentes de um campo tecnológico. Entre os diversos estudos que tratam do universo das patentes, citam-se Breschi e Catalini (2010), Ko *et al.* (2014), Paulo, Ribeiro e Porto (2018), Pinto, Vallone e Honores (2019), Florêncio *et al.* (2020), Bochi e Grácio (2024) e Morandin e Moura (2024).

Entre os conceitos basilares intrínsecos aos estudos de Análise de Rede Social, presentes em Wasserman e Faust (1994), destacam-se aqui os considerados elementares: ator e laço ou ligação relacional. Um ator¹ é uma unidade discreta, individual ou coletiva

1 Na Teoria dos Grafos que fundamenta matematicamente a ARS, um ator é sinônimo de nó. No software Pajek de geração de redes, um ator é sinônimo de vértice.

SUMÁRIO

que corresponde a uma entidade social. Em uma análise de patentes, pode se referir a inventores, país dos inventores, citação a patente, citação a não-patente, classificação, país de depósito, entre outros dados presentes em um documento de patente.

Uma ligação ou laço é a conexão estabelecida entre pares de atores, em função da relação estudada. Entre os tipos de relação analisadas nos estudos de patentes, destacam-se: coinvenção, citação entre patentes (de uma patente para outra), acoplamento tecnológico ou bibliográfico entre as patentes por semelhança às patentes ou literatura por eles citadas, coclassificação, países de codepósito, entre outras.

A partir desses dois conceitos (ator e ligação relacional, tem-se o conceito de *rede social*: um conjunto ou conjuntos finitos de atores e a relação ou relações definidas entre eles.

Na rede ilustrada na figura 7, observam-se 11 atores rotulados pelas letras A a K. Em estudos de patentes, esses 11 atores (nós) podem representar: inventores, países dos inventores, citação (a patente ou literatura científica), classificação, país de depósito, entre outros dados de patente.

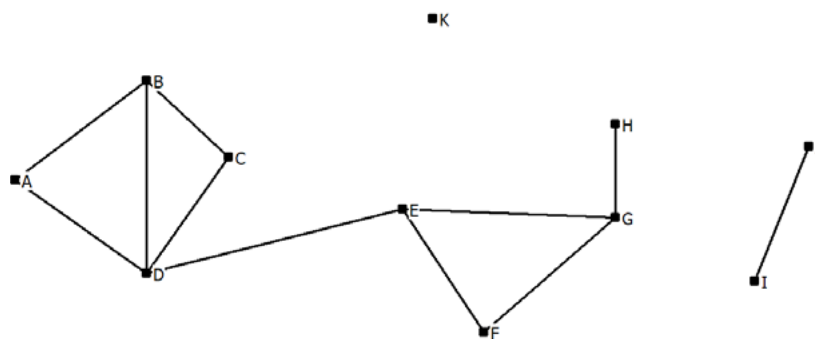
Assim, a título de exemplificação, caso os nós (figura 7) correspondam a 11 inventores, tem-se uma rede de coinventores, com as ligações significando a existência de patentes cocriadas por eles. Nessa situação, K não esteve em parceria com nenhum dos outros 10 inventores e os inventores I e J são coinventores, mas não colaboram na criação de patentes com os demais 9 inventores. Para nós correspondendo a citações em patentes, a figura 7 ilustraria uma rede de cocitação em patentes.

Quanto às ligações, há 11 ligações na rede ilustrativa presente na figura 7, podendo representar as coinvenções, acoplamento bibliográfico ou coclassificação entre os 11 atores, entre outras possibilidades. O fato de o autor k não apresentar ligação com qualquer

SUMÁRIO

outro ator da rede, o denomina *ator isolado*, podendo ser um inventor sem cocriação com os outros analisados.

Figura 7 - Ilustração de rede hipotética composta por 11 atores



Fonte: elaborada pelas autoras, usando Ucinet.

Também o conceito de ator ponte é importante na estrutura de uma rede social. Segundo Wasserman e Faust (1994), um *ator ponte* é aquele responsável pela ligação de dois subgrupos de uma rede. É, portanto, um ator que participa de um subgrupo e interage com um ator de outro subgrupo, isto é, faz a ponte entre dois grupos. Em redes de colaboração científica ou tecnológica, possui um papel importante ao intermediar o fluxo da informação entre dois subgrupos. Na figura 7, os atores D e E são exemplos de atores pontes.

A análise da estrutura de uma rede social utiliza indicadores relativos à rede (densidade) e à posição de cada ator, denominados indicadores de centralidade do ator (grau, proximidade e intermediação). Há diversos estudos na literatura da Ciência da Informação que apresenta seus conceitos, formulação matemática, assim como exemplos de sua aplicação; entre eles, citam-se: Otte e Rousseau (2002), Meyer e Bhattacharya (2004), Boaventura Netto e Jurkiewicz (2009), Matheus e Silva (2009), Lemieux e Ouimet (2004) e Grácio e Oliveira (2020).

SUMÁRIO

Atualmente, há diversos softwares disponíveis para gerar uma rede social e calcular seus indicadores; entre eles, mencionam-se: Ucinet, Pajek, VosViewer e Gephi. Todos são softwares livres, com exceção do Ucinet.

VALORAÇÃO DE PATENTES

Conforme Pita (2010), o método de valoração de patentes consiste na construção de uma função de valor que represente a evolução do valor da patente ao longo dos seus vinte anos de duração, considerando as diversas características desse ativo. A função é construída com base na avaliação dos parâmetros técnicos e econômicos que regem o comportamento das patentes. Para identificar a aderência dos resultados obtidos com evidências encontradas por outros autores é feita uma análise de correlação e teste de hipóteses com os dados bibliométricos das patentes. Os valores são comparados com aqueles apresentados pela literatura específica.

Segundo a Organisation for Economic Co-operation and Development (OCDE, 2009), o valor da patente reflete a diferença entre o valor da invenção com e sem proteção. O valor da tecnologia não é o valor da patente. Pode-se optar por não proteger uma tecnologia. Nesse caso, ela terá um efeito na empresa e no mercado, mas sua exploração será diluída pela concorrência. O preço e o resultado serão diluídos tanto pela pulverização quanto pela comoditização da tecnologia. Mesmo assim, esse intangível tecnológico terá valor para cada uma das empresas no momento em que gerar lucro. A tecnologia protegida a princípio tem dono. Só o seu proprietário a explora, e a patente elimina a concorrência. Os preços de mercado de seus produtos serão maiores, assim como o resultado correspondente. O valor da patente seria, então, o diferencial das duas situações, e não o valor da tecnologia (Adriano; Antunes, 2017).

SUMÁRIO

Cativelli, Pinto e Sanchez (2023), criaram um índice de valoração baseado em três indicadores, cuja nota final será a soma da pontuação, dividida por três. O primeiro indicador considera as famílias de patentes, atribuindo pontos de acordo com o número de países em que uma patente foi depositada, e acrescentando pontos se incluir o Escritório Europeu de Patentes, Estados Unidos e Japão. O depósito via Tratado de Cooperação em Matéria de Patentes (PCT) também é valorizado. O segundo indicador trata das renovações, onde uma patente pontua mais alto a cada renovação paga. Por fim, é avaliado o número de citações, com base em quantas citações foram recebidas, desde a publicação da patente.

PESQUISA DE LEVANTAMENTO

As pesquisas de levantamento caracterizam-se pelo questionamento direto a grupos significativos de indivíduos cujo comportamento se deseja conhecer, sendo associadas à abordagem quantitativa. Quando o levantamento recolhe informações de todos os integrantes do universo pesquisado, tem-se um *censo*. Este tipo de pesquisa proporciona informações gerais acerca das populações, sendo normalmente descritivos.

Em estudos de patentes, este método pode utilizar a técnica de coleta de dados intitulada questionário, para envio a uma comunidade de pesquisadores/inventores, para indagar sobre o uso das patentes e a motivação para formação de parcerias, entre outras questões que possam ser relevantes. Assim, Garcia e Chacon (2008), usaram o questionário para investigar em quais cursos de Biblioteconomia brasileiros a Classificação Internacional de Patentes (CIP) é ensinada. No âmbito de duas universidades brasileiras, Sabino e Garcia (2012) analisaram se a existência de uma política institucional de proteção à PI resulta em maior número de registros de patentes por parte das universidades.

SUMÁRIO

MÉTODOS QUALITATIVOS

Os métodos qualitativos buscam compreender significados, contextos e interpretações relacionadas ao uso e impacto das patentes. Permitem aprofundar a análise a partir da perspectiva social, histórica e epistemológica. Entre os mais utilizados:

ANÁLISE DE DOMÍNIO

Em 1995, Hjørland e Albrechtsen introduziram a Análise de Domínio como uma nova abordagem para a CI. Em 2002, Hjørland propôs 11 maneiras que a CI poderia utilizar para abordar determinado domínio, sendo que o conhecimento combinado dessas abordagens fornece um conhecimento especial aos profissionais da área (Hjørland, 2017).

A abordagem cinco, proposta por Hjørland, trata dos estudos bibliométricos. Em se tratando de patentes, são aplicados indicadores aos dados disponíveis, permitindo descortinar resultados quantitativos. Assim, buscando um pleno aproveitamento do método, esta abordagem deve ser combinada com outras, de forma que uma interpretação qualitativa seja acrescentada.

A pesquisa desenvolvida por Bochi e Grácio (2024) aplicou Análise de Domínio às patentes sobre células-tronco, indexadas na base de dados Derwent Innovation Index. No estudo, as autoras propuseram analisar a contribuição dos métodos de análise de citação univariada e relacional, aplicados a patentes, para identificar e caracterizar domínios científico-tecnológicos. O estudo emprega conjuntamente os métodos de análise de citação relacional: Cocitação e Acoplamento Bibliográfico.

SUMÁRIO

O estudo de Morandin e Moura (2024) aplicou a Análise de Domínio às patentes em Inteligência Artificial depositadas no INPI, combinando a abordagem dos estudos métricos aos estudos históricos e epistemológicos. Esta combinação contextualiza e justifica os resultados encontrados.

ESTUDO DE CASO

O estudo de caso tem se mostrado bastante usado nas pesquisas relacionadas a patentes. Segundo Martins (2008), ele permite um conhecimento aprofundado de determinada realidade social e deve ser precedido por um planejamento detalhado, com a construção de um protocolo que permita reproduzir o estudo. A sistematização de todo o processo inclui notas de observação, registro de opiniões e até mesmo diários.

Pode ser combinado com outros métodos, como a Análise de Conteúdo e utilizar como técnica de coleta de dados, entrevistas. Neste sentido, destaca-se o trabalho de Gonçalves e Tomaél (2013), que verificaram o apoio a inventores em um Núcleo de Inovação Tecnológica (NIT) de uma Universidade do Paraná. Também Santos e Winter (2024) utilizam o estudo de caso para investigar a frequência de pedidos de PI adotados pelo Instituto Federal do Paraná.

REFERÊNCIAS

AALDERING, Lukas Jan; SONG, Chie Hoon. Tracing the technological development trajectory in post-lithium-ion battery technologies: a patent-based approach. **Journal of Cleaner Production**, v. 241, p. 118343, 2019. DOI: <https://doi.org/10.1016/j.jclepro.2019.118343>.

SUMÁRIO

ABBAS, A.; ZHANG, L. SAMEE, U. K. A literature review on the state-of-the-art in patent analysis. **World Patent Information**, [s.l.], v. 37, p. 3-13, 2014. Disponível em: <https://doi.org/10.1016/j.wpi.2013.12.006>. Acesso em: 28 fev. 2025.

ADRIANO, E; ANTUNES, M. T. P. Proposta para mensuração de patentes. **Revista de Administração Contemporânea**, João Pessoa, v. 21, n. 1, 2017. Disponível em: <https://doi.org/10.1590/1982-7849rac201700123> Acesso em: 12 ago. 2025.

BALTAZAR, L. F. *et al.* Patentes como fonte de informação tecnológica para subsídio à pesquisa: uma análise amostral da Universidade Federal do ABC. **Cadernos de Prospecção**, [S. l.], v. 10, n. 4, p. 681, 2017. Disponível em: <https://doi.org/10.9771/cp.v10i4.23208>. Acesso em: 29 set. 2024.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidente da República, [2016]. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 28 fev. 2025.

BOAVENTURA NETTO, P. O.; JURKIEWICZ, S. **Grafos: Introdução e prática**. São Paulo: Editora Blucher, 2009.

BOCHI, F.; GRÁCIO, M.C.C. Characterization of the technological domain of relational citation analyses: A study of stem cell patents. **World Patent Information**, [s.l.], v. 78, 2024. Disponível em: <https://doi.org/10.1016/j.wpi.2024.102302> Acesso em: 28 fev. 2025.

BOCHI, F.; MOURA, A.M.M.; GABRIEL JÚNIOR, R.F. Patentes nos estudos métricos da informação. *In.*: GRÁCIO, M.C.C. *et al.* (Org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária, 2020.

BRESCHI, S.; CATALINI, C. Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. **Research Policy**, [s.l.], v. 39, n. 1, p. 14-26, 2010. Disponível em: <https://doi.org/10.1016/j.respol.2009.11.004> Acesso em: 28 fev. 2025.

BRITO, A. P. D. de ; OZAKI, A. M. **Busca patentária: a chave do sucesso em projetos tecnológicos**. São Paulo: IFSP, 2020 . Disponível em: http://www.ifg.edu.br/attachments/article/1263/INOVA_IFSP_Busca_Patentaria_CONICT20_20.pdf. Acesso em 15 nov. 2023.

CALLON, M.; COURTIAL, J.P.; LAVILLE, F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. **Scientometrics**, Dordrecht, v. 22, p. 155 – 205, 1991. Disponível em: <https://doi.org/10.1007/BF02019280> Acesso em: 28 fev. 2025.

SUMÁRIO

CATIVELLI, A.S.; PINTO, A.L.; LASCURAIN-SANCHEZ, M.L. Construção de um Índice de Valor de Patentes medindo o valor econômico e tecnológico indireto. **Transinformação**, Campinas, v. 35, e220019, 2023. Disponível em: <https://doi.org/10.1590/2318-0889202335e220019> Acesso em: 27 fev. 2025.

CHANG, P. L.; WU, C. C.; LEU, H.-J. Using patent analyses to monitor the technological trends in an emerging field of technology: a case of carbon nanotube field emission display. **Scientometrics**, Dordrecht, v. 82, n. 1, p. 5-19, jan. 2010. Disponível em: <https://doi.org/10.1007/s11192-009-0033-y> Acesso em: 28 fev. 2025.

EGGHE, L.; ROUSSEAU, R. **Introduction to informetrics**: quantitative methods in library, documentation and Information Science. Amsterdam: Elsevier Science Publishers, 1990.

ESPACENET. **Número de pedido**. Espacenet: [Munique], 2017. Disponível em: https://lp.espacenet.com/help?locale=pt_LP&method=handleHelpTopic&topic=applicationnumber

FERREIRA, C. B. T. **O vínculo entre documentos de patentes e a informação obtida em periódicos científicos: estudo aplicado à área câncer de mama**. 2012. 109 f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: <http://ridi.ibict.br/handle/123456789/778> Acesso em: 28 fev. 2025.

FLORÊNCIO, M.N.D.S. *et al.* The sectoral dynamics of the protection of biotechnology in Brazil. **World Patent Information**, [s.l.], v. 62, 2020. Disponível em: <https://doi.org/10.1016/j.wpi.2020.101984> Acesso em: 28 fev. 2025.

GARCIA, J. C. R.; CHACON, F. J. F. O ensino da Classificação Internacional de Patentes (CIP) nos cursos de biblioteconomia brasileiros. **Informação & Informação**, Londrina, v. 13, n. 2, p. 15–33, 2008. Disponível em: <https://doi.org/10.5433/1981-8920.2008v13n2p15> Acesso em: 27 fev. 2025.

GONÇALVES, A. A.; TOMAÉL, M. I. Proteção do conhecimento: apoio a inventores em uma universidade do Paraná. **Informação@Profissões**, [S. l.], v. 2, n. 1, p. 01–21, 2013. Disponível em: <https://doi.org/10.5433/2317-4390.2013v2n1p01> Acesso em: 28 fev. 2025.

GRÁCIO, M. C. C. **Análises relacionais de citação para a identificação de domínios científicos**: uma aplicação no campo dos Estudos Métricos da Informação no Brasil. São Paulo: Universidade Estadual de São Paulo: Oficina Universitária, 2020.

SUMÁRIO

GRÁCIO, M.C.C.; OLIVEIRA, E.F.T. Análise de redes sociais para a visualização do comportamento científico. *In: GRÁCIO, M.C.C. et al. (Org). Tópicos da bibliometria para bibliotecas universitárias*. São Paulo: Editora da UNESP, 2020, p. 164-191.

GUZMÁN SÁNCHEZ, M. V. **Patentometría: herramienta para el análisis de oportunidades tecnológicas**. 130 f. Tese (Doutorado em Gerência de Información Tecnológica)– Facultad de Economía, Universidade de La Habana, Cuba, 1999.

HINZE, S.; SCHMOCH, U. Opening the Black Box. Analytical Approaches and their Impact on the Outcome of Statistical Patent Analysis. *In: Moed H.F., Glanzel W., Schmoch U. (eds). Handbook of Quantitative Science and Technology Research*. Dordrecht: Springer, 2004. Disponível em: https://doi.org/10.1007/1-4020-2755-9_10 Acesso em: 12 ago. 2025.

HJORLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science**, New Jersey, v. 46, n. 6, p. 400-425, 1995. Disponível em: [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-) Acesso em: 28 fev. 2025.

HJORLAND, B. **Domain analysis**. *In: Encyclopedia of Knowledge Organization*. HJORLAND, Birger; GNOLI, Claudio (org.). 2017. Disponível em: https://www.isko.org/cyclo/domain_analysis Acesso em: 07 dez. 2022.

HULLMANN A.; MEYER, M. Publications and patents in nanotechnology. **Scientometrics**, Dordrecht, v. 58, n. 3, p. 507 – 527, 2014. Disponível em: <https://doi.org/10.1023/b:scie.000000687745467a7> Acesso em: 12 ago. 2025.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Manual Básico para Proteção por Patentes de Invenções, Modelos de Utilidade e Certificados de Adição**. Brasília, DF: INPI, 2021. Disponível em: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.gov.br/inpi/pt-br/servicos/patentes/guia-basico/ManualdePatentes20210706.pdf?utm_source. Acesso em: 10 fev. 2025.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Desenho industrial**. Brasília, DF: INPI, 2023. Disponível em: https://www.gov.br/inpi/pt-br/servicos/desenhos-industriais/copy_of_desenhos-industriais. Acesso em: 10 fev. 2025.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Convenção de Paris**. Paris: 1883. Disponível em: <https://www.gov.br/inpi/pt-br/servicos/marcas/arquivos/legislacao/CUP.pdf>

INPI. **Convenção de Berna para a proteção das obras literárias e artísticas**. 1886. Disponível em: <https://www.wipo.int/treaties/en/ip/berne/>.

SUMÁRIO

KO, S.-S. *et al.* Analyzing technology impact networks for R&D planning using patents: combined application of network approaches. **Scientometrics**, Dordrecht, v. 101, n. 1, p. 917 – 936, 2014. Disponível em: <https://doi.org/10.1007/s11192-014-1343-2> Acesso em: 28 fev. 2025.

LEE, L. C.; LEE, Y. Y.; LIAW, Y. C. Bibliometric analysis for development of research strategies in agricultural technology: the case of Taiwan. **Scientometrics**, Dordrecht, v. 93, n. 3, p. 813-830, 2012. Disponível em: <https://doi.org/10.1007/s11192-012-0833-3> Acesso em: 28 fev. 2025.

LEMIEUX, V.; OUIMET, M. **Análise Estrutural das Redes Sociais**. Lisboa: Instituto Piaget, 2004.

LINHARES, F. R. A importância de fazer busca de anterioridade. **Caderno Prospecção**, Salvador, v. 10, n. 4, p. 696-705, out./dez. 2017. Disponível em: <http://dx.doi.org/10.9771/cpv10i4.23235> Acesso em: 28 fev. 2025.

MARTINS, G. A. Estudo de caso: uma reflexão sobre a aplicabilidade em pesquisa no Brasil. **Revista De Contabilidade E Organizações**, [s./l.], v. 2, n. 2, p. 9-18, 2008. Disponível em: <https://doi.org/10.11606/rco.v2i2.34702> Acesso em: 27 fev. 2025.

MATHEUS, R. F.; SILVA, A. B. O. Fundamentação Básica para Análise de Redes Sociais: conceitos, metodologia e modelagem matemática. In: POBLACIÓN, D. A.; MUGNAINI, R.; RAMOS, L. M. S. V. C. (Org.). **Redes Sociais e Colaborativas em informação científica**. São Paulo: Angellara Editora, 2009.

MARCO, A. C. The dynamics of patent citations. **Economics Letters**, [s./l.], v. 94, p. 290-296, 2007. Disponível em: <https://ssrn.com/abstract=949461>. Acesso em: 16 abr. 2018.

MATTOS, A. M.; DIAS, E. W. Análise de cocitação de autores: questões metodológicas. In: Encontro Nacional de Pesquisa em Ciência da Informação, 11, 2010, Rio de Janeiro. **Anais [...]**. Rio de Janeiro: IBICT, 2010.

MEYER, M.; BHATTACHARYA, S. Commonalities and differences between scholarly and technical collaboration: an exploration of co-invention and co-authorship analyses. **Scientometrics**, Amsterdam, 61, n.3, p.443-56, 2004. Disponível em: <https://doi.org/10.1023/B:SCIE.0000045120.04489.80> Acesso em: 28 fev. 2025.

MOEHRLE, M. G. *et al.* Patinformatics as a business process: A guideline through patent research tasks and tools. **World Patent Information**, [s./l.], v. 32, n. 4, p. 291-299, 2010. Disponível em: <https://doi.org/10.1016/j.wpi.2009.11.003>. Acesso em: 12 ago. 2025.

SUMÁRIO

MORANDIN, J. L. P. L.; MOURA, A. M. M. de. Patentes em inteligência artificial no Brasil: caracterização de um domínio tecnológico. **Informação & Informação**, Londrina, v. 29, n. 3, p. 1-26, 2024. Disponível em: <https://doi.org/10.5433/1981-8920.2024v29n3p1> Acesso em: 28 fev. 2025.

MOURA, A. M. M. de *et al.* Fontes de informação em patentes: análise das características das bases Derwent Innovations Index, ORBIT, INPI, Google Patents e PatentScope com base na produção tecnológica da UFRGS. **Folha de Rosto**, v. 5, n. 2, p. 17-27, jan. 2020. Disponível em: <https://lume.ufrgs.br/handle/10183/205631> Acesso em: 28 fev. 2025

MUGNAINI, R.; PAULO de M. J. Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. **Ciência da Informação**, Rio de Janeiro, v. 33, p. 123-131, 2004. Disponível em: <https://doi.org/10.1590/S0100-19652004000200013> Acesso em: 28 fev. 2025.

NGUYEN, A. US20170232781A1. **Device for joining two writing instruments. Depositante: Andre Nguyen.** Data de publicação: 17 ago. 2017. Disponível em: <https://patents.google.com/patent/US20170232781A1/en?q=us20170232781> Acesso em: 28 fev. 2025.

OECD. **OECD Science, Technology and Industry Scoreboard 2009**. Paris: OECD Publishing, 2008. Disponível em: https://www.oecd.org/content/dam/oecd/en/publications/reports/2009/12/oecd-science-technology-and-industry-scoreboard-2009_g1ghb819/sti_scoreboard-2009-en.pdf. Acesso em: 16 ago. 2025.

OTTE, E.; ROUSSEAU, R. Social network analysis: a powerful strategy, also for the information sciences. **Journal of information Science**, [s./], v. 28, n. 6, p. 441-453, 2002. Disponível em: <https://doi.org/10.1177/0165551502028006> Acesso em: 28 fev. 2025.

PAESANI, L. M. **Manual de Propriedade Intelectual**: direito de autor, direito da propriedade industrial, direitos intelectuais, Sui Generis. São Paulo: Atlas, 2012.

PARK, Hyunseok *et al.* A patent intelligence system for strategic technology planning. **Expert Systems with Applications**, v. 40, n. 7, p. 2373-2390, 2013. DOI: <https://doi.org/10.1016/j.eswa.2012.10.073>

PAULO, A.F. *et al.* Mapping countries cooperation networks in photovoltaic technology development based on patent analysis. **Scientometrics**, Dordrecht, v. 117, n. 2, p. 667 – 686, 2018. Disponível em: <https://doi.org/10.1007/s11192-018-2892-6> Acesso em: 28 fev. 2025.

SUMÁRIO

PINTO, P. E.; VALLONE, A.; HONORES, G. The structure of collaboration networks: Findings from three decades of co-invention patents in Chile. **Journal of Informetrics**, [s.l.], v. 13, 2019. Disponível em: <https://doi.org/10.1016/j.joi.2019.100984> Acesso em : 28 fev. 2025.

PITA, A. C. **Análise do valor e valoração de patentes: método e aplicação no setor petroquímico brasileiro**. 2010. Trabalho de Conclusão de Curso (Graduação) – EPUSP, São Paulo, 2010. Disponível em: <https://bdta.abcd.usp.br/directbitstream/45b27fde-66ac-4045-8d47-5799ff20526e/AlbertoCoelhoPita%20TCCPRO10.pdf>. Acesso em: 27 jan. 2025.

PÓVOA, L.; CABELLO, A. **Inteligência Artificial e Desafios para o Sistema de Patentes**. Brasília, DF: Núcleo de Estudos e Pesquisas, 2024 (Texto para Discussão nº 330). Disponível em: <https://www12.senado.leg.br/publicacoes/estudos-legislativos/tipos-de-estudos/textos-para-discussao/td330-a>. Acesso em: 15 de julho de 2024.

PROCACI, T. *et al.* Prospecção Tecnológica: Levantamento de Patentes, Atuação da Academia e Potenciais Inovações em Ambientes de Aprendizagem no Brasil de 2000 a 2015. **iSys-Brazilian Journal of Information Systems**, [s.l.], v. 9, n. 3, p. 69-88, 2016. Disponível em: <https://doi.org/10.5753/isys.2016.316> Acesso em: 28 fev. 2025.

QUINTELLA, C. M. *et al.* Prospecção tecnológica como uma ferramenta aplicada em Ciência e Tecnologia para se chegar à inovação. **Revista Virtual de Química**, Niterói, v. 3, n. 5, 2011. Disponível em: <http://dx.doi.org/10.5935/1984-6835.20110044> Acesso em: 12 ago. 2025.

ROBREDO, J.; CUNHA, M. B. da. Aplicação de técnicas infométricas para identificar a abrangência do léxico básico que caracteriza os processos de indexação e recuperação da informação. **Ciência da informação**, Rio de Janeiro, v. 27, p. 11-27, 1998. Disponível em: <https://doi.org/10.1590/S0100-19651998000100003> Acesso em: 28 fev. 2025.

SABINO, A. G.; GARCÍA, A. Propriedade intelectual em universidades brasileiras: Universidade Federal de Goiás e Universidade de Brasília. **Biblios Revista de Biblioteconomia e Ciência da Informação**, [s.l.], v. 47, p. 82–103, 2012. Disponível em: <https://doi.org/10.5195/biblios.2012.56> Acesso em: 27 fev. 2025.

SANTOS, V. S.; WINTER, E. Apropriabilidade de marcas, patentes, desenhos industriais e indicações geográficas no Instituto Federal do Paraná: um estudo de caso. **P2P E INOVAÇÃO**, Rio de Janeiro, v. 11, n. 1, p. e-7174, 2024. Disponível em: <https://revista.ibict.br/p2p/article/view/7174>. Acesso em: 28 fev. 2025.

SUMÁRIO

URBIZAGÁSTEGUI-ALVARADO, R. Bibliometria brasileira: análise de copalavras.

Transinformação, Campinas, v. 34, p. e220004, 2022. Disponível em: <https://doi.org/10.1590/2318-0889202234e220004> Acesso em: 28 fev. 2025.

VANZ, S. A. S.; CAREGNATO, S. E. Estudos de citação: uma ferramenta para entender a comunicação científica. **Em questão**, Porto Alegre, v. 9, n. 2, p. 295-307, jul./dez. 2003. Disponível em: <https://lume.ufrgs.br/handle/10183/129317> Acesso em: 28 fev. 2025.

WASSERMAN, S.; FAUST, K. **Social network analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Handbook on industrial property information and documentation**. Switzerland: WIPO, 2024. Disponível em: <https://www.wipo.int/export/sites/www/standards/en/pdf/03-09-01.pdf>.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Convenção internacional sobre a proteção dos artistas, intérpretes ou executantes, dos produtores de fonogramas e dos organismos de radiodifusão (convenção de Roma)**. 1961. Disponível em: <https://www.wipo.int/treaties/en/ip/rome/>.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Conferência de Estocolmo para a revisão da convenção da união de Paris para a proteção da propriedade industrial**. 1967. Disponível em: https://www.wipo.int/treaties/en/text.jsp?file_id=283698.

WORLD TRADE ORGANIZATION. **Acordo sobre os aspectos dos direitos de propriedade intelectual relacionados ao comércio (TRIPS)**. 1994. Disponível em: https://www.wto.org/english/docs_e/legal_e/27-trips.pdf.

6

*Thais Dias Medeiros
Nicole Tirello Acquolini
Samile Andréa de Souza Vanz*

METODOLOGIAS PARA ESTUDAR RANKINGS UNIVERSITÁRIOS

DOI: 10.31560/pimentacultural/978-85-7221-475-9.6

SUMÁRIO

COMO CITAR:

MEDEIROS, Thais Dias; ACQUOLINI, Nicole Tirello; VANZ, Samile Andréa de Souza. Metodologias para estudar rankings universitários. *In*: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações**. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

Os *rankings* universitários são uma realidade no contexto do ensino superior no Brasil e no mundo. Diversos autores indicam que, apesar das críticas a seus usos e metodologias, os *rankings* acadêmicos se consolidaram e são utilizados pelas Instituições de Ensino Superior na construção de suas políticas institucionais, como forma de divulgação e ferramenta de marketing (Alves; Dressler; Vanz, 2019; Hazelkorn, 2013). Os *rankings* também são utilizados por estudantes na escolha da instituição para cursar o ensino superior e por docentes e pesquisadores para escolher o local de trabalho, uma vez que uma boa posição nesses *rankings* atribui reconhecimento e prestígio às instituições (De Filippo *et al.*, 2012; Sanz Casado, 2015). Os *rankings* são apontados, ainda, como ferramenta de tomada de decisões de investimento de capital privados e de políticas públicas (Lee, 2020; Martins, 2023; Valmorbida; Cardoso; Ensslin, 2018; Vanz, 2018).

Com a consolidação dos *rankings* universitários surge a preocupação de compreender esses *rankings*, seus usos, suas metodologias, os indicadores utilizados e sua influência no ensino superior. Observa-se um aumento da produção científica sobre a temática de *rankings* acadêmicos no Brasil a partir da década de 2010, com uma presença mais forte e efetiva a partir do ano de 2014, o que indica a emergência dos *rankings* como campo de estudo no país

SUMÁRIO

(Calderón; Pfister; França, 2015; Calderón; França, 2018; Martins, 2023; Wandercil; Calderón; Martins; Ferreira, 2019).

A partir do aumento da produção científica sobre o tema no Brasil é possível identificar uma diversidade de metodologias utilizadas para estudar os *rankings* universitários. Essa diversidade metodológica possibilita uma maior compreensão dos *rankings*, quais aspectos eles avaliam e os indicadores que utilizam. Conhecer as metodologias utilizadas para estudar os *rankings* universitários pode auxiliar na escolha de procedimentos e apontar novos caminhos de pesquisa sobre o assunto. Conforme Martins (2023), a produção acadêmica sobre *rankings* universitários no Brasil se enquadra em duas perspectivas analíticas de avaliação educacional: o paradigma do consenso e o paradigma do conflito.

O paradigma do consenso, segundo Sanders (1984) possui inspiração positivista e funcionalista. Essa perspectiva analítica entende a ordem e o equilíbrio como um objetivo que pode ser alcançado através da comunhão de ideias de grupos sociais e organizações. Essa vertente, nos anos 1960, associava políticas educacionais com questões de trabalho e de economia. As pesquisas sobre *rankings* universitários que se enquadram na abordagem do paradigma do consenso têm uma visão positiva em relação ao uso de *rankings* e indicadores na avaliação da educação. No contexto do paradigma do consenso, então, os *rankings* são vistos como instrumentos utilizados para compreender, avaliar e aprimorar os processos decisórios relacionados à políticas públicas e gestão de Instituições de Ensino Superior (Ferreira, 2013; Martins, 2023; Sanders, 1984).

Já o paradigma do conflito, conforme Sanders (1984), tem base no marxismo e utiliza o modo dialético, e vê o estudo das organizações a partir dos princípios de contradição e de transformação. Para essa perspectiva analítica, as contradições e tensionamentos de grupos sociais e organizações funcionam como fator de transformação. Essa vertente ganha força nas décadas de 1960 e 1970

SUMÁRIO

em contraposição ao paradigma do consenso. Os trabalhos que se enquadram na vertente do paradigma do conflito se baseiam em abordagens críticas ou sócio-críticas, apresentam uma crítica ao pensamento pedagógico liberal e ao uso de indicadores e *rankings* na avaliação da educação, ao processo de mercantilização da educação, e ao produtivismo acadêmico atrelado a ideia de ranqueamento (Ferreira, 2013; Martins, 2023; Sanders, 1984).

Ao estudar as pesquisas sobre *rankings* acadêmicos e governança universitária na literatura científica brasileira, Martins (2023) identificou o interesse da comunidade acadêmica nacional sobre a temática. Em sua pesquisa, encontrou 38 teses e 54 dissertações sobre *rankings* e governança universitária no Catálogo de Teses e Dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e na Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) entre 2014 e 2022. As teses e dissertações sobre *rankings* foram defendidas majoritariamente em instituições públicas e estão distribuídas entre todas as regiões brasileiras, com 32 documentos publicados na região Sudeste, 23 no Sul, 16 no Centro-Oeste, 14 no Nordeste e sete no Norte do país. As teses e dissertações estão vinculadas a nove áreas do conhecimento da CAPES, o que mostra que, apesar da predominância na área da educação e nas grandes áreas de Ciências Humanas e Sociais Aplicadas, os *rankings* universitários são um tema relevante em diversas áreas do conhecimento.

No que diz respeito às abordagens utilizadas nas pesquisas sobre *rankings* acadêmicos e governança universitária, Martins (2023) indica que o paradigma do consenso é predominante nas teses e dissertações (61,66%) e nos artigos (49,73%), especialmente na grande área das Ciências Sociais Aplicadas e nas áreas de Administração Pública e de Empresas, Ciências Contábeis e Turismo. Entretanto, os trabalhos acadêmicos na vertente do paradigma do conflito são predominantes na grande área das Ciências Humanas e na área da Educação.

SUMÁRIO

Martins (2023) também identificou 67 artigos sobre a temática *rankings* e governança universitária publicados no Brasil entre 2017 e 2022. Os periódicos com maior quantidade de publicações sobre *rankings* foram a Avaliação: Revista da Avaliação da Educação Superior, da Universidade Estadual de Campinas (12 artigos), EccoS – Revista Científica da Universidade Nove de Julho e Educar em Revista da Universidade Federal do Paraná (três artigos cada). Assim como acontece com as teses e dissertações, os artigos sobre a temática também estão distribuídos entre todas regiões do país: 36 artigos do Sudeste, 18 do Sul, sete do Centro-Oeste, cinco do Nordeste e um do Norte. Os artigos sobre *rankings*, ao contrário das teses e dissertações que são majoritariamente vinculadas a IES públicas, estão publicados, em sua maioria, em periódicos publicados por instituições privadas.

Ao observar a variedade de abordagens temáticas das pesquisas sobre *rankings* universitários, tanto no paradigma do conflito quanto no paradigma do consenso, percebe-se a preocupação em compreender estes *rankings*, suas características, suas potencialidades e fragilidades bem como os indicadores utilizados de forma ampla e crítica. Os resultados apresentados por Martins (2023) confirmam a consolidação das pesquisas sobre *rankings* universitários e o aumento da sua presença como temática de estudos de pesquisas acadêmicas no país, tendo em vista a importância de conhecer os *rankings* existentes e seus usos na conjuntura da educação superior, compreender suas metodologias e identificar indicadores utilizados.

O presente estudo tem, assim, por objetivo, identificar metodologias utilizadas para estudar *rankings* universitários no Brasil. Para alcançar tal objetivo realizou-se uma revisão de literatura dos textos brasileiros sobre *rankings* universitários, a partir dos primeiros textos publicados, na década de 1990. Após leitura, realizou-se uma organização das pesquisas a partir dos métodos utilizados, categorizando os textos da seguinte maneira: a) pesquisas teóricas sobre *rankings* universitários, que se baseiam em análises históricas

SUMÁRIO

e contextuais sobre os *rankings*; b) pesquisas teóricas e empíricas que abrangem análise quantitativa dos indicadores e metodologias utilizadas pelos *rankings* para a avaliação das Instituições de Ensino Superior; e c) pesquisas empíricas que analisam o desempenho das Instituições de Ensino Superior, cursos e disciplinas específicas nos *rankings* universitários. A seguir apresenta-se uma revisão da literatura organizada a partir destas três categorias, reunindo exemplos de estudos e aplicações.

PESQUISAS TEÓRICAS ACERCA DA HISTÓRIA E CONTEXTO DOS *RANKINGS*

Os estudos teóricos fazem parte das primeiras pesquisas sobre *rankings* no Brasil, realizados majoritariamente por pesquisadores da área da Educação. O primeiro documento publicado que se tem registro foi escrito por Schwartzman em 1995 na revista 'Ensaio: Avaliação e Políticas Públicas em Educação' e intitulado 'Dificuldades e possibilidades de se construir um *ranking* para as universidades brasileiras,' no qual sugere a elaboração de um *ranking* das universidades brasileiras, baseado em três indicadores: o Índice de Qualificação do Corpo Docente (IQCD), utilizado como parâmetro para aferir a produção acadêmica; a avaliação de Cursos de Graduação realizada na época através do Guia do Estudante, publicado pela Editora Abril; e a avaliação da pós-graduação, executada pela CAPES (Calderón; Poltronieri; Borges, 2011).

O interesse dos pesquisadores da Educação foi despertado pelas inúmeras transformações que o sistema de educação superior no Brasil passou desde a metade da década de 1990, especialmente com a criação de universidades privadas com fins lucrativos, as chamadas "universidades mercantis." Um dos pesquisadores que

SUMÁRIO

discute o tema é Calderón (2000), em um estudo configurado como ensaio teórico que, apesar de não detalhar a metodologia empregada, demonstra um caráter bibliográfico ao apresentar o estado da arte sobre o tema.

De acordo com Martins (2023), as instituições privadas ampliaram o acesso ao ensino superior, contrastando com o modelo público, que mantinha políticas restritivas focadas em pesquisa e vagas limitadas. Em meio a esse novo cenário, o governo brasileiro desenvolveu políticas de avaliação para assegurar padrões de qualidade, surgindo mecanismos como o Programa de Avaliação da Reforma Universitária (PARU) em 1983, seguido pelo Programa de Avaliação Institucional das Universidades Brasileiras (PAIUB) em 1993. Em sua abrangente tese sobre *rankings* acadêmicos e governança universitária, Martins (2023) realizou um mapeamento de estudos sobre o tema, utilizando uma abordagem multimetodológica que combina técnicas quantitativas, de natureza bibliográfica e bibliométrica, com métodos qualitativos baseados na análise de conteúdo, e no estado da arte, destacando os avanços na produção do conhecimento sobre a temática no Brasil.

O Exame Nacional de Cursos (ENC), popularmente chamado de Provão, foi implementado em 1995 com o intuito de classificar as instituições de ensino superior com base no desempenho dos discentes. Essa iniciativa marcou uma mudança na cultura da avaliação no ensino superior brasileiro, onde o objetivo passou a ser também a fiscalização e regulação estatal sobre a qualidade do ensino (Martins, 2023). O Provão é considerado o primeiro *ranking* surgido no cenário nacional, o que justifica inúmeros estudos sobre sua metodologia e as críticas atribuídas (Calderón; Poltronieri; Borges, 2011).

Em 2004, num contexto que “[...] cada vez mais se tornava abrangente e complexo [...]” (Martins, 2023, p. 32), o Sistema Nacional de Avaliação da Educação Superior (SINAES) expandiu seus métodos de avaliação, incorporando fatores como ensino, pesquisa, extensão,

SUMÁRIO

responsabilidade social, gestão da instituição e corpo docente, que passaram a ser analisados por meio dos dados do Exame Nacional de Desempenho de Estudantes (ENADE), avaliações institucionais e de cursos (Brasil, 2022; Martins, 2023).

Sobrinho (2008), docente e pesquisador da área da Educação, analisou o SINAES por meio de uma compilação de discussões apresentadas na mesa-redonda “Qualidade, pertinência e responsabilidade social da Educação Superior”, no Seminário Internacional “Educação Superior: Avaliação e Tendências na América Latina e no Caribe” em 2008. A avaliação objetivista e produtivista, segundo Sobrinho (2008), vincula-se às ideologias do individualismo e da competitividade, próprias de uma sociedade cujo valor central é o mercantilismo. Diversos autores fazem duras críticas a esses critérios e procedimentos avaliativos, os quais são vistos como excessivamente controladores e técnicos, e, por conseguinte, sem importância social para as comunidades educativas. Isso se deve, principalmente, ao fato de não conseguirem atender aos conceitos de pertinência local e relevância social, ou seja, à chamada “qualidade social”.

No entanto, não seria adequado sustentar que todos os autores que questionam os métodos de avaliação adotados pelo governo, bem como o uso dos resultados dessas avaliações, negligenciam a importância da avaliação educacional. Ao contrário, certos autores destacam a necessidade das avaliações como um meio para assegurar a qualidade da educação em seus distintos níveis e modalidades de ensino, assim como aponta Sobrinho (2008, p. 822) “Os processos de avaliação e acreditação são importantíssimos, sobretudo em face dos crescentes fenômenos de mercantilização, diversificação, fragmentação [e] internacionalização [...]”, ou seja, as avaliações servem como um recurso para garantir a eficácia e a direção de possíveis reformas no contexto atual. E, enquanto processo de construção de sentidos e reflexão sobre valores, possui também uma extensa capacidade educativa.

SUMÁRIO

Neste mesmo estudo, Sobrinho (2008) também é enfático ao associar o SINAES ao fenômeno de ranqueamento, originado no ambiente empresarial norte-americano, no cenário do *benchmarking* – mecanismo de análise de mercado que realiza comparações entre uma empresa e suas concorrentes dentro do mesmo setor –, visando atender aos interesses de gestores empresariais e formuladores de políticas, e não aos interesses dos educadores de modo geral. Além disso, Soares (2022, p. 35) comenta que nesse fenômeno “Aliado ao sistema disciplinar do neoliberalismo, há o emprego de recompensas, normalmente associadas à produtividade”.

Outras pesquisas foram publicadas com a intenção de investigar e mapear o SINAES e o ENADE. Lacerda, Ferri e Duarte (2016), a título de exemplo, examinaram as práticas de operacionalização do SINAES, desenvolvido sob o critério da eficiência institucional em seu produto final, o aluno concluinte. Os autores discutem as modalidades do sistema: o exame nacional dos cursos, a avaliação externa dos cursos de graduação e o processo de autoavaliação, e analisam a lógica do desempenho por meio de índices criados pelo sistema. O estudo aponta que o documento original do SINAES previa um sistema distinto daquele que efetivamente ocorreu na prática: uma forma incipiente de um modelo de *accountability* – expressão que evoca a ideia de transparência nos processos – baseado em exame de desempenho. Como conclusão, Lacerda, Ferri e Duarte (2016) propõem que se reexamine o papel do ENADE e seus respectivos índices derivados, Conceito Preliminar de Cursos (CPC) e Índice Geral de Cursos (IGC), no contexto do SINAES, a fim de que eles sejam utilizados como indicadores iniciais do processo avaliativo e não o ponto final deste processo.

Embora o governo tenha inicialmente justificado o SINAES como um modelo de avaliação formativa e emancipatória, múltiplas ações de aprimoramento do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) foram implementadas, como a criação do CPC e do IGC que segundo Martins (2023)

SUMÁRIO

consolidaram o sistema em um formato próximo ao do Provão, numa perspectiva técnica-burocrática “[...] concorrencial e legitimadora de valores neoliberais, gerando listas classificatórias e estabelecendo ranqueamento das Instituições de Ensino Superior (IES)”, assumindo um viés “[...] meritocrático, classificatório e promotor de ranqueamentos” (Martins, 2023, p. 33).

O CPC e o IGC, por sua vez, também foram objeto de inúmeros estudos teóricos, como o de Ikuta (2016) que discute a concepção, aplicação e mudanças metodológicas do CPC a partir de uma revisão bibliográfica e documental sobre o índice. Além da apresentação do SINAES e da descrição dos processos que conduziram à criação e aplicação do índice, a autora traz as revisões metodológicas do CPC ao longo dos anos. Ikuta (2016) revela que a primeira revisão aconteceu em 2008, quando os componentes da fórmula passaram de apenas três para oito, além da mudança da ponderação do peso desses componentes, com o aumento do peso atribuído à titulação docente. Já na segunda revisão, em 2011, a autora destaca que foi dispensada do cálculo do PCP a Nota dos Ingressantes do ENADE (NI). Em 2013, em nova alteração, ampliaram-se os itens sobre a percepção discente a partir da utilização do Questionário do Estudante do ENADE (Ikuta, 2016). A autora conclui salientando a importância de debater e problematizar os pontos fortes e fracos do CPC. Entre as fragilidades, Ikuta (2016) elenca a falta de avaliação presencial dos cursos, as incertezas sobre consistência técnica geradas pelas diversas mudanças de cálculo do índice, a homogeneização das notas em relação às especificidades de cursos e de instituições, além da falta de transparência do processo como um todo para a sociedade.

Já o IGC é analisado por Bittencourt, Casartelli e Rodrigues (2009) que buscam compreender e discutir o indicador, além de identificar estratégias que as IES podem utilizar para melhorar esse índice. Para os autores, a divulgação IGC pela imprensa teve como efeito transformar o indicador em instrumento de marketing institucional que influencia na formação de opinião pública. Após decompor os

SUMÁRIO

conceitos considerados pelo índice, os autores indicam que o Índice Geral de Cursos deveria levar em conta os resultados da Avaliação Externa. Uma crítica ao índice é não ponderar as especificidades de cursos e instituições (Bittencourt; Casartelli; Rodrigues, 2009).

Antes do Estado consolidar seu papel como avaliador e regulador da educação superior, os *rankings* acadêmicos começavam a se estabelecer no Brasil. No setor privado, desde 1980, já existiam iniciativas de ranqueamento das Instituições de Ensino Superior (IES). Um exemplo disso foi o *ranking* Melhores Faculdades, criado em 1981 e publicado anualmente por 19 anos seguintes pela revista masculina Playboy, configurando-se numa ação de classificação de IES no país (Martins, 2023).

Nesse sentido, o Brasil é o único país no âmbito do ensino superior de língua portuguesa a contar com *rankings* promovidos tanto pela iniciativa pública quanto pela privada, acompanhando assim um movimento global (Martins, 2023). Um exemplo adicional de *ranking* oriundo da iniciativa privada é o *Ranking Universitário Folha* (RUF), elaborado pelo jornal de grande circulação *Folha de S. Paulo* entre os anos de 2012 e 2019. O RUF foi objeto de análise de Soares (2022) em sua dissertação intitulada *Ranking Universitário da Folha*, que, a partir de pesquisa qualitativa e documental comparativa, explorou o *ranking* buscando examinar e definir a ideia de excelência universitária utilizada e as possíveis consequências de sua utilização. O autor constatou que o conceito de excelência universitária privilegia indicadores de pesquisa, principalmente aqueles referentes à produtividade e impacto, e descarta totalmente (e de forma surpreendente, dada a missão institucional das próprias universidades brasileiras) a avaliação das atividades de extensão das universidades. Além disso, foi possível examinar a correlação entre este específico modo político de funcionamento da estatística e da construção de índices com a discussão existente na literatura sobre racionalidade neoliberal. Soares (2022) reforça a percepção de que os *rankings*, longe de ser apenas uma métrica e números neutros,

SUMÁRIO

são peças centrais nos processos de fabricação de um contexto competitivo entre as universidades brasileiras.

Outra proposta de avaliação de IES realizada pelo setor privado foi o *Guia do Estudante*, publicado anualmente pela Editora Abril (mencionado anteriormente), entre 1984 e 2018. Porém, segundo Righetti (2016), essa avaliação não ranqueava, apenas classificava os cursos superiores em categorias de “bons”, “muito bons” ou “excelentes”. Para conduzir sua tese sobre política científica e tecnológica, Sabine Righetti apresentou uma revisão da literatura e um estudo de caso voltado a *rankings* universitários na gestão da USP (Righetti, 2016), que será detalhada na penúltima seção deste capítulo. Mais recentemente, segundo Martins (2023) surgiram outras iniciativas privadas de ranqueamento, como o *Ranking* Nacional de Universidades Empreendedoras, organizado pela Confederação Brasileira de Empresas Juniores e o EAD *Ranking*, primeiro *ranking* de educação a distância do Brasil, que já avaliou 200 cursos (Martins, 2023). Além dessas ações, fazendo uma busca rápida do Google, é possível identificar, atualmente, diferentes propostas não formais de *rankings* que avaliam a educação superior brasileira.

Em âmbito global, a primeira lista classificatória de universidades que se tem registro na literatura científica foi elaborada em 1983 pelo jornal *U.S. News*. Desde então, a concepção de *rankings* acadêmicos influenciou alguns países, como China, Alemanha e outros países europeus a adotarem essa prática (Righetti, 2016). De acordo Hazelkorn (2013), que investiga diretamente o início e a evolução dos *rankings* acadêmicos, foi com a expansão da globalização — marcada pela transição para um mercado mundial integrado de bens e serviços — que produtos, serviços e até distintos aspectos da vida cotidiana passaram a ser aferidos por meio de indicadores, em que as vantagens competitivas e comparativas podem ser determinantes. Ainda assim, para a autora, os *rankings* representam apenas

SUMÁRIO

um dos muitos instrumentos disponíveis para avaliar o desempenho institucional ou aumentar a transparência. Sua relevância, em grande parte, decorre da simplicidade, uma vez que seus resultados são facilmente compreendidos.

Corroborando e complementando, Vanz (2018) afirma que os *rankings* universitários internacionais têm se tornado cada vez mais relevantes no contexto global, pois simplificam e condensam uma grande quantidade de dados sobre universidades de distintos países. Embora essa simplificação seja sua principal vantagem, a redução de informações provenientes de diferentes contextos econômicos, sociais e culturais pode enviesar os resultados. Essa lógica resulta, dentre outras razões, de que os *rankings* universitários são um dos resultados inevitáveis da popularização da educação superior (Martins, 2023).

Vieira e Lima (2016) realizaram a análise de conteúdo em uma pesquisa bibliográfica e documental, com vistas a colaborar para uma discussão sobre as consequências da utilização dos *rankings* globais na busca pelo selo 'Universidade Classe Mundial' (UCM). Foram examinados o *Academic Ranking of World Universities* (ARWU), o *Times Higher Education* (THE), o *Quacquarelli Symonds World University Rankings – The Top 200 World Universities* (QS). As autoras observaram que a determinação de conquistar o selo UCM contribuiu para valorizar métodos de avaliação comparativos e imprimir um caráter estratégico à gestão acadêmica das universidades. Esse efeito colateral reforça o processo de internacionalização vigente e aproxima as instituições educacionais da lógica organizacional que regula as corporações. Vieira e Lima (2016) concluíram que, em meio a um ambiente competitivo, o selo UCM contribui para legitimar a competitividade da universidade, favorecer a atração de estudantes internacionais (pagantes) e a captação de recursos financeiros capazes de promover um ciclo virtuoso pelo fortalecimento da pesquisa.

SUMÁRIO

De acordo Herting (2021), os *rankings* universitários globais são filhos do neoliberalismo turbo-globalizado (TGNL). Conceito que segundo o autor foi crucial para um expressivo crescimento econômico em nível global, mas que falhou em garantir uma distribuição justa da riqueza gerada – as desigualdades dentro das sociedades aumentaram e regiões inteiras do mundo são negligenciadas – e se mostraram incapazes de enfrentar e mitigar grandes crises globais como as mudanças climáticas. Como efeito, um novo *Zeitgeist* – termo alemão que, em sua tradução literal, significa “espírito do tempo”, mas que, segundo Krause (2019), refere-se a padrões culturais específicos de um período – está emergindo, com foco em um crescimento econômico direcionado para a qualidade, sustentabilidade, equidade, justiça social e equilíbrio entre o global e o local. Nesse contexto, as universidades no pós-TGNL, assim como os métodos para avaliar sua qualidade, assumem uma nova face, devido a uma reviravolta ideológica em percurso no ensino superior internacional (Herting, 2021).

Herting (2021), com base em um estudo empírico sobre *rankings* conduzido por ele em 2016, apresenta, em uma nova pesquisa sobre o tema, uma análise teórica que destaca cinco aspectos fundamentais para que os *rankings* universitários se alinhem ao *Zeitgeist*. Em suma, esses aspectos envolvem uma abordagem de avaliação multidimensional, pautada na justiça social e na democratização, evitando critérios arbitrários, como a quantidade de laureados com o Prêmio Nobel ao longo do tempo (recorrente em alguns *rankings* internacionais). O modelo proposto considera indicadores de desempenho regional, valorizando o conhecimento local e as metodologias de pesquisa específicas de cada cultura como contribuições significativas para a produção científica. Além disso, confere igual relevância ao ensino, pesquisa e extensão, bem como avalia o impacto social e a devolutiva do conhecimento gerado pelas instituições acadêmicas (Herting, 2021). Em função desses aspectos, a próxima seção explora diversas pesquisas sobre os indicadores e metodologias empregadas nos *rankings*.

SUMÁRIO

PESQUISAS SOBRE INDICADORES
E METODOLOGIAS UTILIZADAS
PELOS *RANKINGS*

Entre as pesquisas sobre *rankings* realizadas no Brasil estão as que têm seu foco na análise dos indicadores e metodologias utilizadas por diversos *rankings* para a avaliação das Instituições de Ensino Superior. Esses estudos utilizam metodologias teóricas e empíricas para analisar os indicadores e metodologias nos quais se baseiam diferentes *rankings* regionais, nacionais e/ou internacionais.

Calderón, França e Gonçalves (2017), por exemplo, realizaram um estudo de caráter descritivo-analítico e comparativo, por meio de pesquisa bibliográfica e documental de *rankings* de abrangência nacional ou regional realizados por jornais ibero-americanos de grande circulação: *50 Carreras – Los Mejores Centros Universitarios*, do jornal *El Mundo*; *Ranking de Calidad de las Universidades Chilenas*, do *El Mercurio*; *Ranking Universitario Folha – RUF*, do *Folha de São Paulo*; *Las Mejores Universidades*, do *Reforma* e *Mejores Universidades*, do *El Universal*. Os autores propõem o “Grid de tipologias dos *rankings* acadêmicos”, no qual ampliam a ideia proposta alguns anos antes por Andrade (2011) para a definição de tipologias dos *rankings* a partir dos indicadores que utilizam. O Grid categoriza os *rankings* segundo sua tipologia: os *rankings* puros (divididos em *rankings* objetivos e subjetivos) e *rankings* híbridos.

No trabalho original, Andrade (2011) esmiúça os indicadores utilizados pelos *rankings* governamentais ENEM e ENADE, argumentando acerca da importância de conhecermos que tipo de informação cada *ranking* fornece. A partir de diversas análises quantitativas e qualitativas, o autor conclui que o MEC tinha, à época, uma política distinta para o *ranking* no Ensino Superior em comparação ao do Ensino Médio. Andrade (2011) sugere diversas alternativas para

SUMÁRIO

melhoria das fórmulas e indicadores utilizados. Adicionalmente, o *ranking* privado *Guia do Estudante* também é analisado pelo estudo.

Valmorbida, Cardoso e Ensslin (2018), por sua vez, realizaram pesquisa bibliográfica em combinação com pesquisa-ação para descobrir se há consonância entre os indicadores e os objetivos dos *rankings* *Academic Ranking of World Universities* (ARWU), *Times Higher Education Supplement – World University Ranking* (THES), *Ranking Higher Education Evaluation and Accreditation Council of Taiwan* (HEEACT), *Leiden Ranking – Centre for Science and Technology Studies* (CWTS), *Ranking Web ou Webometrics*, *SCImago Institutions Ranking* (SIR), *Quacquarelli Symonds World University Rankings* (QS) e *America's Best College Review – Us News and World Report*. A partir dessa metodologia, Valmorbida, Cardoso e Ensslin (2018) analisaram os objetivos e indicadores desses *rankings*, quem participa de sua construção, para quem se destinam, se existe congruência entre objetivos e indicadores e se os indicadores são gerenciáveis pelas instituições avaliadas.

Já Vanz (2018) realizou revisão de literatura para discutir as características dos *rankings* *Shangai*, THE, QS e *Leiden*, categorizando os *rankings* baseados em indicadores qualitativos – avaliação de opinião (*surveys*), no qual se enquadram os *rankings* THE e QS, e aqueles baseados em indicadores quantitativos, como os *rankings* *Shangai* e *Leiden*. Segundo Vanz (2018) os *rankings* baseados em indicadores qualitativos – avaliação de opinião, são alvo de críticas pela subjetividade desse tipo de análise quando se trata de *rankings*. Já os *rankings* que utilizam indicadores quantitativos, apesar de também apresentarem desvantagens, são considerados mais objetivos por terem como base dados sólidos como produção científica e citação.

Guimarães (2020) trabalhou os mesmos *rankings* a partir de outra vertente metodológica. O autor analisou os *rankings* THE, *Shangai*, *Leiden* e QS por meio de sua natureza classificatória

SUMÁRIO

utilizando os Sistemas de Organização do Conhecimento, proveniente da Ciência da Informação, para analisar os aspectos e vieses culturais desses *rankings*. O autor organizou os elementos analíticos dos *rankings* segundo notação, critérios de atualização, facetas (indicadores e subindicadores) e modalidades de edições. O trabalho aponta que esse conhecimento organizado sobre *rankings* é “[...] subsídio fundamental à geração de novo conhecimento seja no que tange ao planejamento e à autoavaliação institucional seja em ações voltadas à divulgação e à visibilidade das universidades” (Guimarães, 2020, p. 17).

Soares (2022) por sua vez, realizou um levantamento sobre a metodologia e indicadores utilizados pelo RUF, descrevendo que este possuía, em sua primeira edição, quatro indicadores: pesquisa (composto por diversos indicadores baseados principalmente em produção científica), inovação (baseado em indicadores de patente), mercado e ensino (baseado em pesquisa de reputação). Em 2013 houve inclusão de novos indicadores e alteração nos pesos atribuídos para os indicadores já utilizados. Em 2014 a metodologia do RUF se solidificou em relação a pesos atribuídos e indicadores: pesquisa (peso 42), ensino (32), mercado (18), internacionalização (quatro) e inovação (quatro), com mudanças posteriores em relação aos componentes dos indicadores.

Leal, Stallivieri e Moraes (2018) analisaram os principais *rankings* acadêmicos, com foco nos indicadores relacionados à internacionalização da educação superior. Em uma pesquisa de natureza qualitativa e descritiva, desenvolvida por meio de consultas bibliográficas e documentais, os autores avaliaram seis *rankings*: *Academic Rankings of World Universities* (ARWU); *Q&S World University Rankings* (Q&S); *Times Higher Education World University Rankings* (THE); *Ranking Universitário Folha* (RUF); *Ranking Guia do Estudante* (GE) e *Conceito Preliminar de Curso* (CPC) / Índice Geral de Cursos (IGC).

SUMÁRIO

Manganote, Schulz e Cruz (2016) realizaram testes empíricos analisando o efeito dos artigos da *High Energy Physics Large Collaboration* nos *rankings Scimago* e THE. Os autores argumentaram que é necessário considerar a forma como a pesquisa é desenvolvida para avaliar se o indicador citação pode ser usado para definir *rankings* individuais ou institucionais. Os resultados apresentados pelos autores comprovaram o efeito positivo dos artigos da área de Física de Altas Energias nos resultados de *output* e citação, tanto em nível institucional quanto de país. Segundo os autores, a participação em grandes colaborações (*big science*) pode distorcer resultados de universidades com pequeno volume de *output*, conforme resultados da pesquisa empírica empreendida. Manganote, Schulz e Cruz (2016) concluem que a Física de Altas Energias influencia positivamente o desempenho das universidades nos *Ranking THE*, apesar da metodologia divulgada à época pelo *ranking* afirmar o contrário. Os autores sugerem a necessidade de uma discussão sobre metodologias e intervalos de tempo distintos a serem considerados pelos *rankings*.

Tendo em vista a importância do tema avaliação e *rankings* universitários, o Conselho de Reitores (Cruesp) das universidades estaduais paulistas desenvolveu o projeto “Indicadores de Desempenho nas Universidades Estaduais Paulistas”, que resultou na publicação de dois livros intitulados “Repensar a Universidade”, lançados em 2018 e 2019 (Marcovitch, 2018; 2019). Conforme texto de apresentação das obras, foi unânime a decisão por empreender um estudo conjunto sobre políticas a serem adotadas e definir qual a base digital ideal para aferição, registro e difusão do seu desempenho acadêmico, mantidas, naturalmente, a vocação distinta das três universidades estaduais paulistas, quais sejam, USP, UNESP e UNICAMP. As três instituições dispõem de unidades estatísticas e cada uma delas produz o seu anuário de indicadores. Os capítulos dos livros Repensar a Universidade são de autoria de vários estudiosos e pesquisadores, e apresentam ensaios teóricos sobre avaliação e sobre os indicadores utilizados pelos *rankings*, e algumas análises empíricas do desempenho das universidades em questão.

SUMÁRIO

Prado (2022) realizou uma revisão das dimensões e indicadores do *ranking* multidimensional europeu *U-Multirank*. Para o autor o caráter multidimensional desse *ranking* permite uma avaliação que leva em conta diversos indicadores, o que permite uma visão geral das instituições para seus futuros alunos. Prado (2022) traz que no *U-Multirank* não há uma classificação entre melhores e piores instituições, e que seu desempenho é apresentado a partir de seus pontos fortes e fracos. Nesse *ranking* são utilizados dados das próprias instituições, bancos de dados bibliométricos e de patentes, além de banco de dados nacionais e questionários (Prado, 2022). Para sua análise, Prado (2022) fez um estudo sobre as instituições acadêmicas espanholas e depois utilizou os resultados encontrados para identificar pontos das dimensões e indicadores desse *ranking* que possam ter utilidade para a pós-graduação brasileira. Para o autor, essa nova metodologia proposta pode trazer mudanças significativas nas notas dos programas de pós-graduação no país e, conclui, ainda, que esse método pode ser utilizado também para avaliar a graduação (Prado, 2022).

Souza *et al.* (2019) propõem um índice composto que utiliza indicadores bibliométricos para analisar a atividade científica de universidades federais do Brasil entre 2003 e 2015. Para isso os autores utilizam como fonte de informação o Sistema e-MEC, o Sistema de Informações Georreferenciadas (GeoCAPES), a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), o Instituto Nacional da Propriedade Industrial do Brasil (INPI) e a Base de dados da *Web of Science*. Souza *et al.* (2019) calculam, então, a partir das informações coletadas, o Índice Composto, utilizando variáveis de acordo com duas classificações: (1) variáveis dependentes do tamanho da instituição, que inclui, total de documentos; número de teses; total de citações recebidas; total de documentos em periódicos do primeiro quartil (Q1) e do primeiro decil (D1); número de documentos com colaboração internacional; número de documentos publicados nos periódicos *Nature* e *Science*; e (2) variáveis independentes do

SUMÁRIO

tamanho da instituição, que considera, média de citações por documento; média de citações de professor-pesquisador. Os autores concluem que o Índice Composto pode servir como método alternativo para as instituições do Brasil que não estejam ranqueadas. Recomendam, ainda, a utilização de fontes locais para coleta de informações (Souza *et al.*, 2019).

Assim, após a exposição de algumas pesquisas que discutiram indicadores e metodologias utilizadas pelos *rankings*, a seção seguinte abordará os estudos que analisam o desempenho das universidades e seus respectivos cursos e disciplinas.

PESQUISAS EMPÍRICAS SOBRE O DESEMPENHO DAS UNIVERSIDADES E CURSOS/DISCIPLINAS ESPECÍFICAS

Foram observados muitos estudos sobre o desempenho das universidades e cursos/disciplinas ou *rankings* específicos. Estes estudos se referem a um período específico ou abrangem uma série longitudinal de dados, com intuito de apresentar a análise evolutiva das instituições brasileiras. Diferentemente dos estudos teóricos, observam-se aqui diversos estudos da bibliometria, a maioria proveniente da área da Ciência da Informação.

A literatura científica sobre *rankings* universitários publicada mais recentemente aborda o tema sob vários aspectos e é praticamente unânime ao afirmar que *rankings* universitários são um caminho sem volta e, por isso, precisam ser estudados. A tese de doutorado de Righetti (2016) inicia com uma análise teórica sobre a representação do ensino superior na sociedade atual, até chegar ao debate sobre avaliação universitária por meio de *rankings*.

SUMÁRIO

universitários globais e nacionais. Deste ponto em diante, a tese utiliza abordagem empírica a partir da análise dos indicadores e resultados obtidos pela USP. Conforme a autora, a USP, apesar de afirmar que não utiliza as listagens na gestão, monitora os *rankings* e os utiliza para marketing.

Santos e Noronha (2016) analisaram os indicadores e o desempenho longitudinal das universidades brasileiras em três dos principais *rankings* internacionais: ARWU, THE e QS no período de 2003 a 2014. A análise revelou que as universidades brasileiras se destacam por sua pontuação no indicador volume de publicações e em alguns casos também no indicador reputação entre acadêmicos e empregadores. Pontuações mais baixas foram obtidas nos indicadores que medem impacto por meio de publicações altamente citadas, indicador que costuma ter alto peso na maioria dos *rankings*. Os resultados obtidos pelas universidades brasileiras foram comparados com os obtidos pelas universidades mais bem colocadas em cada um dos *rankings*. As autoras concluíram que as classificações internacionais são, de modo geral, homogeneizadoras, de maneira que muitas particularidades, nuances e bolsões de excelência em áreas específicas não são capturados pelos *rankings* internacionais quando as instituições são avaliadas como um todo.

Vanz *et al.* (2018a) analisaram os principais indicadores do *Leiden Ranking* e o posicionamento das 16 universidades brasileiras presentes na edição 2016. Os resultados mostraram um panorama bastante diverso no que diz respeito ao volume de produção e colaboração internacional entre as universidades brasileiras presentes no *Leiden Ranking* 2016. A USP destacou-se das demais universidades brasileiras quanto ao volume de produção científica. No entanto, no que diz respeito ao impacto medido através de citações, observou-se um padrão no desempenho obtido. O desempenho por áreas do conhecimento também foi detalhado no estudo, revelando o protagonismo da USP nas cinco áreas estudadas (*Life, Biomedical, Mathematics, Physical* e *Social Sciences*). Uma das principais

SUMÁRIO

conclusões do estudo diz respeito aos periódicos utilizados pelas universidades para comunicação de seus resultados de pesquisa. Entre as 146 revistas brasileiras presentes na lista de periódicos do *Leiden Ranking* 2016, apenas 25 (17%) fazem parte do *core journals*. Àquela época, os autores sugeriram que a ampliação do volume de produção científica em *core journals* poderia se converter em estratégia básica para melhoria do posicionamento das universidades brasileiras no *Leiden Ranking*.

Vanz *et al.* (2018b) analisaram o desempenho de 29 universidades brasileiras classificadas nos *rankings* *Leiden*, *Shanghai*, QS e THE. Os resultados obtidos permitiram concluir que o desempenho das universidades brasileiras é muito bom, especialmente no que diz respeito ao volume de produção científica, no entanto, o impacto, medido pelas citações recebidas, ainda constitui um grande desafio. O volume de artigos publicados é responsável, junto com os artigos publicados nas revistas *Nature* e *Science*, pela presença de seis universidades brasileiras (USP, UNICAMP, UFMG, UFRJ, UFRGS e UNESP) no *Ranking de Shanghai*. O volume de publicações nas revistas *Nature* e *Science*, assim como a ampliação da colaboração internacional, são estratégias que podem ampliar o impacto e abrir caminho para o alcance de melhores posições nestes *rankings*. Outras estratégias propostas pelos autores se referem ao financiamento das publicações na *Nature* e *Science*, tendo em vista que ambas operam com taxas de publicação (*Article Processing Charge* - APC), e o controle e correção dos nomes dos pesquisadores e de suas instituições de vínculo.

Alves, Dressler e Vanz (2019) realizaram um estudo com o objetivo de identificar a preocupação com os *rankings* por parte de universidades brasileiras a partir de dois aspectos: a formalização do tema em seus Planos de Desenvolvimento Institucional (PDIs) e o aparecimento do tema em notícias publicadas em seus sites institucionais. Para isso, os autores aplicaram análise documental dos PDIs de 29 universidades brasileiras e notícias publicadas, revelando,

SUMÁRIO

naquele momento, pouco interesse pelos *rankings* tendo em vista que o tema pouco apareceu nos PDIs analisados. Em relação às notícias, os autores observaram de maneira preliminar que, por mais que exista a divulgação do desempenho obtido nos *rankings*, aparentemente apenas as universidades privadas utilizam os *rankings* como uma forma de marketing.

Morandin, Silva e Vanz (2020), realizaram análise das metodologias utilizadas por dois *rankings* distintos, o *ranking* internacional e multidimensional *U-Multirank* e o nacional e unidimensional *Ranking Universitário Folha* (RUF), buscando expandir a perspectiva de compreensão acerca do tema e o aprofundamento da discussão acerca destes *rankings* no Brasil. Adicionalmente, realizou-se a análise de desempenho das universidades brasileiras nos anos 2019/2020. As autoras fundamentaram-se em uma pesquisa bibliográfica e documental, com coleta de dados nos sites dos respectivos *rankings*. Os resultados mostraram que as instituições de ensino brasileiras são grandes e generalistas, o que dificulta a visibilidade internacional. Porém, o resultado do *ranking* nacional apresentado mostra-se similar ao internacional, na comparação entre as instituições. A análise da metodologia utilizada pelos *rankings* indica os pontos negativos e positivos que influenciam na avaliação das universidades. A visualização dos resultados no "*U-Multirank*" através de conceitos e gráficos multidimensionais permite a identificação dos pontos fortes das instituições além de aspectos que podem ser aprimorados. As instituições brasileiras têm poucos indicadores conceituados, o que decorre, possivelmente, do não preenchimento/envio dos questionários. Os indicadores qualitativos deste *ranking* poderiam ser a mola propulsora das IES nacionais.

Duarte, Alves e Vanz (2023) utilizaram dois métodos distintos para avaliar o desempenho das universidades brasileiras em quatro *rankings* acadêmicos internacionais e posterior divulgação deste desempenho em notícias na mídia. Os autores buscaram os resultados obtidos por 54 universidades brasileiras nos *rankings* ARWU,

SUMÁRIO

Leiden Ranking, QS e THE, edições 2020 e 2021, mostrando que o número de universidades nacionais presentes nos *rankings* quase dobrou em relação a estudo anterior, e este aumento traz maior visibilidade internacional para as universidades brasileiras, o que pode acarretar maiores investimentos e parcerias. O maior volume de universidades ranqueadas também resultou em maior presença dos Estados brasileiros, visto que as 54 universidades estão sediadas em 21 Estados diferentes. A segunda parte do estudo recuperou 582 notícias publicadas por estas universidades entre os anos de 2010 e 2020, através de uma pesquisa documental das notícias no Google. As notícias foram analisadas de forma quantitativa, para identificar o número total de publicações encontradas, bem como o número de notícias publicadas por ano. Em seguida foi feita a leitura completa e categorização das notícias a partir das temáticas, seguindo procedimentos de análise de conteúdo.

A análise de conteúdo aplicada às notícias revelou que, à medida em que o número de universidades brasileiras classificadas pelos *rankings* internacionais cresce, fica claro que a importância dada para estes *rankings* cresce também. A análise das notícias publicadas revela que as instituições brasileiras vêm demonstrando maior interesse neste assunto, seja por questões de reconhecimento, parcerias ou investimentos. Por mais que a maioria das notícias encontradas por Duarte, Alves e Vanz (2023) seja de divulgação dos resultados, os autores se surpreenderam ao encontrar um número razoável de notícias sobre medidas tomadas pelas universidades para melhorar o desempenho nos *rankings*. Isso evidencia a preocupação das universidades brasileiras com os resultados obtidos nos *rankings* internacionais.

Cabello e colaboradores (2019) estudaram a classificação de instituições brasileiras de ensino superior no QS *World University Ranking* e no *The Times Higher Education* (THE), avaliando as restrições metodológicas e como essas restrições afetaram a classificação das universidades do país. Os autores mostraram que entre

SUMÁRIO

2016 e 2019, 38 universidades brasileiras foram classificadas no QS e no THE, em sua versão global. Devido a restrições metodológicas, a classificação das instituições brasileiras assumiu a característica de patamares, variando de três a cinco patamares. Os autores consideram como patamar uma posição em torno da qual um conjunto de IES brasileiras se aglomera, provavelmente porque o indicador não é sensível o suficiente para diferenciá-las frente às instituições internacionais. Cabello e colaboradores (2019) concluíram que a classificação de uma instituição em um patamar mais alto depende de sua proximidade com um mercado de trabalho dinâmico e um orçamento maior, indicando a dificuldade dos *rankings* de diferenciar de fato qualidade além dessas duas dimensões.

Vanz e Docampo (2022a; 2022b) realizaram estudo exploratório com artigos brasileiros indexados no *InCites* entre 2010-2019 e recalcularam pesos para artigos em colaboração com quatro países de língua inglesa - Austrália, Canadá, Estados Unidos e Reino Unido. Os dados do *InCites* foram cotejados com o desempenho das universidades brasileiras presentes nas 54 disciplinas do *Global Ranking of Academic Subjects* (ARWUGRAS) edição 2021. Os autores mostraram que o Brasil figura em 39 das 54 disciplinas do ARWUGRAS, com a presença de 49 diferentes universidades brasileiras. Ao se retirar a colaboração internacional com países de língua inglesa, observou-se alteração no posicionamento das universidades brasileiras na lista de classificação. Os testes estatísticos demonstraram vantagem significativa para a colaboração com estes países, considerando todas as disciplinas juntas, e que a porcentagem de publicações em periódicos do primeiro quartil dobrou. Quando testadas de forma específica por grandes áreas, observou-se que não há ganho significativo para Engenharias e Ciências da Vida, mas há resultados significativos para a vantagem apresentada pelas Ciências Naturais e Ciências da Saúde. A pesquisa conclui que há vantagem no posicionamento das IES nacionais no ARWUGRAS quando existe colaboração com autores provenientes da Austrália, Canadá, Estados Unidos e Reino Unido.

SUMÁRIO

Wandercil e colaboradores (2024), analisaram comparativamente a evolução histórica que as universidades brasileiras tiveram nos três principais *rankings* nacionais: o *Ranking Universitário Folha*, o *Guia do Estudante* e o Índice Geral de Cursos (IGC), a partir de uma análise bibliográfica e documental de fontes secundárias de informação. Os resultados obtidos pelos autores mostraram que existe um conjunto de instituições que se situam sempre num determinado intervalo de posições, independentemente do *ranking*, da metodologia utilizada ou do ano em questão. A maior parte destas instituições são públicas e apenas duas são privadas, mas denominacionais, filantrópicas e sem fins lucrativos. Conforme Wandercil *et al.* (2024), dentre o universo de 203 universidades brasileiras, existe um seletor grupo de excelência formado por apenas 13 universidades, segundo os critérios dos três *rankings* analisados.

Após a exposição de algumas pesquisas empíricas sobre o desempenho das universidades e seus respectivos cursos e disciplinas, são apresentadas, a seguir, as considerações finais fundamentadas no que foi abordado até o momento.

CONSIDERAÇÕES FINAIS

Este capítulo apresenta uma revisão de literatura acerca dos estudos sobre *rankings* universitários destacando publicações de autores brasileiros, com o intuito de mapear as metodologias utilizadas. No Brasil, o estudo sobre *rankings* foi iniciado na década de 1990 pelos pesquisadores da Educação. Essas investigações eram de cunho teórico e abordavam a avaliação universitária, discutindo as dificuldades e as possibilidades de criar um *ranking* para as universidades brasileiras naquela época. A pesquisa bibliográfica e documental predomina entre os primeiros estudos, com destaque para estudos de caso.

SUMÁRIO

Até hoje, os estudos sobre esse tema são desenvolvidos majoritariamente por pesquisadores da Educação. Após alguns anos, a partir dos anos 2010, e acompanhando o surgimento do *Ranking* de Shanghai e a consequente popularização de vários outros *rankings*, observamos o despertar de outros grupos de pesquisadores provenientes de áreas distintas. Podem-se mencionar os pesquisadores provenientes de escolas de Administração, Economia, Sociologia e Ciência da Informação. Percebe-se, então, o aumento do número de pesquisas e publicações sobre a temática, assim como a diversificação e metodologias relacionadas aos *rankings* acadêmicos. Estes grupos propõem estudos teóricos que analisam os indicadores e metodologias utilizados pelos *rankings*, mas o foco é direcionado para análise métrica das fórmulas e fontes de informação utilizadas pelos *rankings*. Os estudos empíricos das instituições ou de áreas do conhecimento também são comuns entre estes pesquisadores, que replicam a metodologia dos *rankings* e apresentam estudos preditivos.

Além disso, no Brasil, observa-se produção científica sobre a temática em todas as regiões, tanto em instituições de ensino superior públicas quanto privadas. O interesse crescente também é demonstrado pelo volume de teses e dissertações sobre o tema, além de artigos. As revistas científicas onde são publicados estes trabalhos também denotam interesse crescente, tendo em vista que são revistas vinculadas a áreas do conhecimento distintas, como Educação, Economia, Administração e Ciência da Informação.

Esse aumento de teses, dissertações e artigos em distintas áreas sobre o assunto demonstra o interesse da comunidade acadêmica em compreender a dinâmica desses *rankings*, suas metodologias e o impacto que eles exercem nas políticas educacionais e na gestão das Instituições de Ensino Superior. A partir da revisão da literatura, foi possível identificar uma diversidade de abordagens metodológicas que abrangem estudos teóricos, quantitativos e empíricos, oferecendo uma compreensão mais ampla e crítica sobre os *rankings* e suas implicações. Assim, a presença de dois paradigmas

SUMÁRIO

analíticos – o do consenso e o do conflito – revela a complexidade e as múltiplas perspectivas sobre os *rankings* universitários. Portanto, a presente pesquisa destaca a expressividade de uma análise crítica e multidisciplinar sobre os *rankings* universitários, considerando suas metodologias e indicadores.

Apesar do aumento e da diversificação, a revisão de estudos brasileiros sobre *rankings* universitários revela um conjunto pequeno de trabalhos, se comparado aos estudos internacionais. Ainda há muito para se desenvolver em termos de quantidade, e também, no que diz respeito à aplicação de metodologias distintas. Espera-se que o presente estudo sirva de apoio na escolha metodológica para pesquisadores que pretendem estudar os *rankings* universitários, bem como na identificação de lacunas de estudos na temática. Observa-se a existência de estudos internacionais bastante focados na reprodução das metodologias informadas pelos *rankings*, e simulação a partir de conjuntos de dados referentes às universidades obtidos em bases de dados como a *Web of Science* e SCOPUS. O desenvolvimento de conhecimento estatístico e matemático, bem como o acesso a fontes de dados como *Incites* e *SciVal* podem apoiar o desenvolvimento deste tipo de estudo no Brasil.

REFERÊNCIAS

ALVES, Artur Basílio Venturella; DRESSLER, Carolina Georg; VANZ, Samile Andrea de Souza. Rankings acadêmicos nas universidades públicas brasileiras: presença no planejamento estratégico e nas notícias institucionais. *In*: CALDERÓN, Adolfo Ignacio; WANDERCIL, Marco; MARTINS, Edivaldo Camarotti. (org.). **Rankings acadêmicos e governança universitária no espaço do ensino superior de língua portuguesa: Angola, Cabo Verde, Macau, Moçambique, Portugal e Brasil**. Brasília: Anpae, 2019. p. 164-181. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/204098/001108646.pdf?sequence=1>. Acesso em: 10 out. 2024.

SUMÁRIO

ANDRADE, Eduardo de Carvalho. Rankings em Educação: tipos, problemas, informações e mudanças: análise dos principais rankings oficiais brasileiros. **Estudos Econômicos**, São Paulo, v. 41, n. 2, p. 323-343, abr./jun. 2011. DOI: <https://doi.org/10.1590/S0101-41612011000200005>. Acesso em: 24 nov. 2024.

BITTENCOURT, Hélio Radke; CASARTELLI, Alam de Oliveira; RODRIGUES, Alziro César de Moraes. Sobre o índice Geral de Cursos (IGC). **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, v. 14, n. 3, p. 667-682, nov. 2009. Disponível em: http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S1414-40772009000300008&lng=pt&nrm=iso. Acesso em: 12 fev. 2025.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **O que é o SINAES?** 2022. Disponível em: <http://inep80anos.inep.gov.br/inep80anos/presente/sinaes/125>. Acesso em: 26 out. 2024.

CABELLO, Andrea Felipe; IMBROISI, Denise; FALQUETO, Junia Maria Zandonade; FERREIRA, Guilherme Viana; ARRUDA, June Alvez de. Rankings Universitários Internacionais: evidências de vieses geográficos e orçamentários para intuições brasileiras. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, v. 24, n. 03, p. 637-657, set./nov. 2019. DOI: <https://doi.org/10.1590/S1414-40772019000300005>. Acesso em: 12 out. 2024.

CALDERÓN, Adolfo Ignacio. Universidades Mercantis: a institucionalização do mercado universitário em questão. **São Paulo em Perspectiva**, São Paulo, v. 14, n. 1, p. 61-72, 2000. DOI: <https://doi.org/10.1590/S0102-88392000000100007>.

CALDERÓN, Adolfo Ignacio; POLTRONIERI, Heloisa; BORGES, Regilson Maciel. Os rankings na educação superior brasileira: políticas de governo ou de Estado? **Ensaio**: avaliação e Políticas Públicas em Educação, Rio de Janeiro, v. 19, n. 73, p. 813-826, out./dez. 2011. DOI: <https://doi.org/10.1590/S0104-40362011000500005>. Acesso em: 19 nov. 2024.

CALDERÓN, Adolfo Ignacio; FRANÇA, Carlos Marshal. Rankings acadêmicos na educação superior: tendências da literatura ibero-americana. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, v. 23, n. 2, p. 448-466, jul. 2018. DOI: <https://doi.org/10.1590/S1414-40772018000200010>. Acesso em: 15 out. 2024.

SUMÁRIO

CALDERÓN, Adolfo Ignacio; FRANÇA, Carlos Marshal; GONÇALVES, Armando. Tendências dos rankings acadêmicos de abrangência nacional de países dos espaço Ibero-Americano: os rankings dos jornais El Mundo (Espanha), El Mercurio (Chile), Folha de São Paulo (Brasil), Reforma (México) e El Universal (México). **EccoS - Revista Científica**, São Paulo, n. 44, p. 117-141, set./dez. 2017. DOI: <https://doi.org/10.5585/eccos.n44.7943>. Acesso em: 15 out. 2024.

CALDERÓN, Adolfo Ignacio; PFISTER, Mariana; FRANÇA, Carlos Marshal. Rankings Acadêmicos na Educação Superior Brasileira: a emergência de um Campo de Estudos (1995-2013). **Roteiro**, Joaçaba, v. 40, n. 1, p. 31-50, jan./jun. 2015. DOI: <http://dx.doi.org/10.18593/rv40i1.6440>.

DE FILIPPO, Daniela; CASANI, Fernando; GARCÍA-ZORITA, Carlos; EFRAÍN-GARCÍA, Preiddy; SANZ-CASADO, Elías. Visibility in international rankings. Strategies for en racing the competitiveness of Spanish universities. **Scientometrics**, Amsterdam, v. 93, n. 3, p. 949-966, 2012. DOI: <https://doi.org/10.1007/s11192-012-0749-y>. Acesso em: 02 mar. 2025.

DUARTE, Monique Finn; ALVES, Artur Basilio Venturella; VANZ, Samile Andrea de Souza Vanz. As universidades brasileiras nos rankings universitários internacionais: desempenho e divulgação. In: CALDERÓN, Adolfo Ignacio; WANDERCIL, Marco; FONSECA, Dora Maria Ramos; VANZ, Samile Andrea de Souza (orgs.). **A construção de universidades de classe mundial e rankings acadêmicos no espaço do Ensino Superior de Língua Portuguesa e em outras realidades do mundo**. Brasília: Anpae, 2023. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/270418/001185677.pdf?sequence=1>

FERREIRA, Tânia Aparecida. **Exame Nacional de Cursos**: um balanço da produção científica brasileira, com foco nos impactos nas IES e suas estratégias de aprimoramento institucional (1999-2010). Dissertação (mestrado em Educação) – Programa de Pós-Graduação em Educação, Centro de Ciências Humanas e Sociais Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, 2013.

GUIMARÃES, José Augusto Chaves. Os rankings universitários como Sistemas de Organização do Conhecimento: considerações teóricas. **Informação & Sociedade**, João Pessoa, v. 30, n. 4, p. 1-23, 2020. DOI: <https://doi.org/10.22478/ufpb.1809-4783.2020v30n4.57187>. Acesso em: 26 out. 2024.

HAZELKORN, Ellen. How Rankings are Reshaping Higher Education. In: CLIMENT, Vicent; MICHAVILA, Francesc; RIPOLLÉS, María. (eds.). **Los rankings universitarios, Mitos y Realidades**. Madrid: Técnos, 2013.

SUMÁRIO

HERTIG, Hans Peter. Benchmarking and zeitgeist: university ranking in post-turbo-globalized neoliberalism (Post-TGNL). *In*: HAZELKORN, Ellen; MIHUT, Georgiana.

Research Handbook on University Rankings: theory, methodology, influence and impact. Cheltenham: Edward Elgar Publishing, 2021. p. 93-104. Disponível em: <https://www.elgaronline.com/edcollbook/book/9781788974981/9781788974981.xml>. Acesso em: 05 nov. 2024.

IKUTA, Camila Yuri Santana. Sobre o Conceito Preliminar de Curso: concepção, aplicação e mudanças metodológicas. **Estudos em Avaliação Educacional**, São Paulo, v. 27, n. 66, p. 938-969, set./dez. 2016. DOI:<https://doi.org/10.18222/aeae.v27i66.4039>. Acesso em: 02 mar. 2025.

KRAUSE, Monika. What is Zeitgeist? Examining period-specific cultural patterns. **Poetics**, [s.l.], v. 76, p. 101352, 2019. DOI: <http://dx.doi.org/10.1016/j.poetic.2019.02.003>. Acesso em: 03 dez. 2024.

LACERDA, Leo Lynce Valle de; FERRI, Cássia; DUARTE, Blaise Keniel da Cruz. SINAES: avaliação, accountability e desempenho. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, v. 21, n. 3, p. 975-992, nov. 2016. DOI: <http://dx.doi.org/10.1590/S1414-40772016000300015>. Acesso em: 03 dez. 2024.

LEAL, Fernanda Geremias; STALLIVIERI, Luciane; MORAES, Mário César Barreto. Indicadores de internacionalização: o que os rankings acadêmicos medem? **Revista Internacional de Educação Superior**, Campinas, v. 4, n. 1, p. 52-73, 2018. DOI: <https://doi.org/10.22348/riesup.v4i1.8650638>. Acesso em: 03 out. 2024.

LEE, Rosemary Theodoro. **Doutorados profissionais em educação física**: tendências em Universidades de Classe Mundial (UCM). 2020. Dissertação (Mestrado em Educação) – Programa de Pós-graduação em Educação, Centro de Ciências Humanas e Sociais Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, 2020.

MANGANOTE, Edmilson J. T.; SCHULZ, Peter A.; CRUZ, Carlos Henrique de Brito. Effect of high energy physics large collaborations on higher education institutions citations and rankings. **Scientometrics**, Amsterdam, v. 109, n. 2, p. 813-826, 2016. DOI: <https://doi.org/10.1007/s11192-016-2048-5>.

MARCOVITCH, Jacques (org.). **Repensar a universidade**: desempenho acadêmico e comparações internacionais. São Paulo: Com-Arte; Fapesp, 2018. 256 p. (Projeto Métricas). DOI: <https://doi.org/10.11606/9788571661868>. Acesso em: 03 mar. 2025.

SUMÁRIO

MARCOVITCH, Jacques (org.). **Repensar a Universidade II: Impactos para a Sociedade**. São Paulo: Com-Arte; Fapesp, 2019. 328 p. (Projeto Métricas). DOI: <https://doi.org/10.11606/9788571661967>. Acesso em: 03 mar. 2025.

MARTINS, Edivaldo Cesar Camarotti. **Rankings acadêmicos e governança universitária**: mapeamento e tendências da literatura científica brasileira. Tese (Doutorado em Educação) – Programa de Pós-Graduação em Educação, Centro de Ciências Humanas e Sociais Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, São Paulo, 2023.

MORANDIN, Janaina Lais Pacheco Lara; SILVA, Natalia Rodrigues da; VANZ, Samile Andrea de Souza. O desempenho das universidades brasileiras no U-Multirank e Ranking Universitário Folha. **Ciência da Informação em Revista**, Maceió, v. 7, n. 2, 2020, p. 116-136. DOI:10.28998/cirev.2020v7n2h

RIGHETTI, Sabine. **Qual é a melhor?**: origem, indicadores, limitações e impactos dos rankings universitários. 2016. Tese (Doutorado em política científica e tecnológica) - Instituto de Geociências, Universidade Estadual de Campinas, Campinas, 2016. Disponível em: <http://repositorio.unicamp.br/handle/REPOSIP/321911>. Acesso em: 05 nov. 2024.

RIGHETTI, Sabine. Avaliar para comparar: os rankings britânico e chinês no ensino superior global. In: MARCOVITCH, Jacques. (org.). **Repensar a universidade**: desempenho acadêmico e comparações internacionais. São Paulo: Com-Arte; Fapesp, 2018. p. 45–62. (Projeto Métricas). DOI: <https://doi.org/10.11606/9788571661868>.

PRADO, Antonio Fernando Bertachini de Almeida. Avaliação Acadêmica Multidimensional com o uso do “U-Multirank”. **Avaliação**: Revista da Avaliação do Ensino Superior, Campinas, v. 27, n. 1, p. 159-182, mar. 2022. DOI: <https://doi.org/10.1590/S1414-40772022000100009>. Acesso em: 12 fev. 2025.

SANDER, Benno. **Consenso e Conflito**: Perspectivas Analíticas na Pedagogia e na Administração da Educação. São Paulo: Livraria Pioneira; Rio de Janeiro: Universidade Federal Fluminense, 1984. (Biblioteca Pioneira de Ciências Sociais).

SANTOS, Solange Maria; NORONHA, Daisy Pires. O desempenho das universidades brasileiras em rankings internacionais. **Em Questão**, Porto Alegre, v. 22, n. 2, p. 186–219, 2016. DOI: <https://doi.org/10.19132/1808-5245222.186-219>. Acesso em: 19 nov. 2024.

SANZ-CASADO, Elias. (coord.). **Guía de buenas prácticas para la participación de las universidades españolas en los rankings internacionales**. Madrid: Ministerio de Educación, Cultura y Deporte. 2015. 101 p.

SUMÁRIO

SOBRINHO, José Dias. Qualidade, avaliação: do SINAES a índices. **Avaliação** Revista da Avaliação da Educação Superior, Campinas, v. 13, n. 3, p. 817-825, 2008. DOI: <https://doi.org/10.1590/S1414-40772008000300011>.

SOARES, Arthur de Ávila. **Rankings universitários em perspectiva: o Ranking Universitário Folha e a Construção da excelência**. Dissertação (Mestrado em Sociologia) – Faculdade de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Sociologia, Universidade Federal de Minas Gerais, Minas Gerais, 2022.

SOUZA, Cláudia Daniele de; SUÁREZ-BALSEIRO, Carlos; MAURA-SARDÓ, Mariano; DE FILIPPO, Daniela; SANZ-CASADO, Elías. Aplicación de um índice compuesto para el análisis de la actividad científica de las universidades brasileñas. **Ciência da Informação**, Brasília, v. 48, n. 3, p. 81-99, set./dez. 2019. DOI: <https://doi.org/10.18225/ci.inf.v48i3.4625>. Acesso em: 03 mar. 2025.

VALMORBIDA, Sandra Mara lesbik; CARDOSO, Thuine Lopes; ENSSLIN, Sandra Rolim. Rankings universitários mundiais: análise da congruência entre objetivos e indicadores. **Revista Gestão Universitária na América Latina- GUAL**, Florianópolis, v. 11, n. 1, 2018. DOI: <https://doi.org/10.5007/1983-4535.2018v11n1p229>. Acesso em: 03 dez. 2024.

VANZ, Samile Andréa de Souza. O que medem os rankings universitários internacionais? Apontamentos teóricos, indicadores e características. **Informação & Sociedade: Estudos**, João Pessoa, v. 28, n. 2, p. 83-92, maio/ago., 2018. Disponível em: <https://periodicos.ufpb.br/index.php/ies/article/view/38383>. Acesso em: 18 out. 2024.

VANZ, Samile Andrea de Souza; PANDIELLA DOMINIQUE, Andres; LASCURAIN SÁNCHEZ, María Luisa; SANZ-CASADO, Elías. O desempenho das universidades brasileiras no Leiden Ranking. In: BENETTI, Marcia; BALDISSERA, Rudimar. **Pesquisa e perspectivas de comunicação e informação**. Porto Alegre: Sulina, 2018a. p. 276-292.

VANZ, Samile Andrea de Souza; PANDIELLA DOMINIQUE, Andres; LASCURAIN SÁNCHEZ, María Luisa; SANZ-CASADO, Elías. Rankings universitários internacionais e o desafio para as universidades brasileiras. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 23, n. 53, p. 39-51, 2018b. DOI: <https://doi.org/10.5007/1518-2924.2018v23n53p39>. Acesso em: 19 nov. 2024.

VANZ, Samile Andrea de Souza; DOCAMPO, Domingo. Colaboração brasileira com autores da Austrália, Canadá, Estados Unidos e Reino Unido e o desempenho das universidades no ranking ARWUGRAS. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 27, n. 1, p. 1-21, 2022. DOI: <https://doi.org/10.5007/1518-2924.2022.e84306>. Acesso em: 19 nov. 2024.

SUMÁRIO

VANZ, Samile Andréa de Souza; DOCAMPO, Domingo. The influence of International Scientific Collaboration with English-speaking countries on the research performance of Brazilian Academic Institutions. **Journal of Scientometric Research**, New Delhi, v. 11, n. 3, p. 358-370, sep./dec. 2022. DOI: <https://doi.org/10.5530/jscires.11.3.39>. Acesso em: 03 mar. 2025.

VIEIRA, Rosilene Carla; LIMA, Manolita Correia. A busca pelo selo “universidade classe do mundo” e os rankings acadêmicos globais pioneiros. **Interfaces Brasil/Canadá**, Canoas, v. 16, n. 1, p. 97-116, 2016. DOI: <https://doi.org/10.15210/interfaces.v16i1.7704>. Acesso em: 03 mar. 2024.

WANDERCIL, Marco; CALDERÓN, Adolfo Ignacio; MARTINS, Edivaldo Camarotti; FERREIRA, Tânia Aparecida. **Rankings acadêmicos: um panorama da Educação superior brasileira**. In: CALDERÓN, Adolfo Ignacio; WANDERCIL, Marco; MARTINS, Edivaldo Camarotti (orgs). *Rankings acadêmicos e governança universitária no espaço do ensino superior de língua portuguesa: Angola, Cabo Verde, Macau, Moçambique, Portugal e Brasil*. Brasília: Anpae, 2019.

WANDERCIL, Marco; GANGA-CONTRERAS, Francisco; IGNACIO CALDERÓN, Adolfo; ABELLO-ROMERO, Juan; ARAYA-CASTILLO, Luis; SAN MARTÍN, Walter Sáez. Sistema educativo brasileño: universidades de excelencia a la luz de los rankings académicos nacionales*. **Práxis Educativa**, Ponta Grossa, v. 19, e22663, p. 1-23, 2024. DOI: <https://doi.org/10.5212/praxeduc.v19.22663.071>. Acesso em: 22 fev. 2025.

7

*Sônia Elisa Caregnato
Pedro Henrique da Silva Rodrigues
Carolina Carvalho Sena
Geise Ribeiro da Silva*

MÉTODOS DE PESQUISA EM COMUNICAÇÃO CIENTÍFICA:

**ESTUDOS DE IMPACTO E VISIBILIDADE
POR MEIO DE CITAÇÕES ABERTAS**

SUMÁRIO

COMO CITAR:

CAREGNATO, Sônia Elisa; RODRIGUES, Pedro Henrique da Silva; SENA, Carolina Carvalho; SILVA, Geise Ribeiro da. Métodos de pesquisa em comunicação científica: estudos de impacto e visibilidade por meio de citações abertas. /n: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações**. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

Citar as fontes que fundamentam um trabalho científico é uma prática consolidada entre pesquisadores, que são introduzidos a ela desde os primeiros estágios da formação acadêmica. Trata-se de uma norma essencial para a comunicação da ciência com qualidade, pois visa reconhecer os trabalhos anteriores e estabelecer vínculos entre o documento que cita e os documentos citados.

Essa prática social permitiu, ao longo do tempo, e especialmente a partir da criação do *Citation Index*, que se consolidasse o uso do conjunto das citações de uma determinada área do conhecimento para três finalidades fundamentais: recuperar trabalhos similares, acompanhar o desenvolvimento histórico da área e gerar indicadores para a avaliação da atividade científica.

O uso das citações como elementos de avaliação não foi algo que surgiu com Eugene Garfield, quando, na metade da década de 1960, propôs um índice de citações, o *Citation Index*, cujo objetivo era oferecer uma nova abordagem para o problema de recuperação da informação na literatura científica (Garfield, 1955). Sua perspectiva era a da recuperação da informação, no entanto sua criação impulsionou a análise de citações.

SUMÁRIO

Como observam Aksnes, Langfeldt e Wouters (2009), proliferaram, nas últimas décadas, indicadores bibliométricos voltados à avaliação da pesquisa e das políticas de pesquisa. E é essa aplicação da contagem de citações nas avaliações (desde revistas e pesquisadores até instituições e países) e a criação do Fator de Impacto, decorrente do *Citation Index*, que causa a polêmica que se observa hoje acerca das citações. A menção a citações costuma ser, para o interlocutor leigo, imediatamente associada ao Fator de Impacto e a outras métricas que as utilizam, mas que não as resumem.

Uma entre tantas críticas aos indicadores derivados de citações se dirige à seletividade das bases de dados utilizadas para seu cálculo, que priorizam a indexação de periódicos predominantemente publicados por grandes editoras de nações desenvolvidas e em língua inglesa. A força deste argumento se dilui com o surgimento de várias novas fontes de dados bibliográficos, que são alternativas aos *Citation Indexes* da *Web of Science*.

Superando esses vieses, estudos mostram que citar é um processo cognitivo e social complexo, que permite estabelecer relações semânticas entre artigos e que demonstra a ligação entre os textos e os autores. Melhor dizendo, citações são marcas do conhecimento anterior deixadas em textos mais recentes (Cronin, 1984) ou “associação de ideias” (Garfield, 1955, p. 108).

No final das contas, citações em textos científicos refletem a forma como as publicações são recebidas pelas comunidades científicas, na medida em que são utilizadas em textos posteriores. Ou seja, podem ser entendidas como indicadores de recepção.

As citações geram visibilidade na Academia, representando uma dimensão do impacto da pesquisa na ciência e, algumas vezes, na tecnologia (por meio de citações de artigos em documentos de patentes ou de vice-versa). O impacto mais amplo na sociedade, no entanto, precisa ser dimensionado de outras formas, o que já vem sendo buscado por alguns pesquisadores, inclusive os que utilizam as altmetrias.

SUMÁRIO

Partindo da premissa de que os estudos de citação têm aplicações que vão além da avaliação, este capítulo propõe apresentar elementos conceituais para sua compreensão, além de discutir as principais fontes de informação disponíveis e ferramentas para coleta de dados, especialmente considerando que a Ciência Aberta também demanda a utilização de citações abertas.

FUNDAMENTAÇÃO PARA OS ESTUDOS DE CITAÇÃO

Os estudos de citação têm como objeto as menções feitas em textos científicos a outras publicações, geralmente detalhadas nas listas de referências. Esses estudos são aplicados às diferentes áreas do conhecimento, especialmente por meio de métodos quantitativos. A facilidade de gerar estatísticas permite que se estude as citações aos artigos de um periódico, de um autor, de um grupo de pesquisa, de uma instituição ou de um país.

A decisão de citar é uma escolha do autor, que define quem citar, quando e como incorporar o trabalho de terceiros ao seu próprio texto. No entanto, essa prática só se concretiza plenamente em um contexto social, o que faz da citação um dos principais elementos no sistema de comunicação científica.

Entre as discussões centrais sobre o fenômeno da citação, destaca-se aquela que considera os diversos motivos para citar como argumento para questionar a citação enquanto indicador de impacto, uma vez que os trabalhos podem ser citados por múltiplas razões, inclusive por críticas ou desqualificações, e não apenas por mérito ou reconhecimento intelectual.

SUMÁRIO

Silveira, Caregnato e Bufrem (2014) propõem uma classificação para as razões de citação, estruturada em duas dimensões principais: a conceitual e a social. A primeira se refere a citações feitas para apontar aspectos teóricos, metodológicos e de opinião, ou seja, quando são utilizadas para apresentação de definições, explicações teóricas ou de pressupostos, para atualização, para sustentar argumentos, apresentar exemplos ou reforçar determinados tópicos ou ainda criticar e negar abordagens. A dimensão social, por sua vez, compreende as razões relacionadas à tradição científica, como o reconhecimento de autores pioneiros, a menção a trabalhos históricos, a retificação de trabalhos anteriores, a disputa por prioridade de ideias e a contraposição a posicionamentos.

Além das razões de citação, alguns pesquisadores da comunicação científica buscaram interpretar as perspectivas teóricas que embasam os estudos de citação (Nicolaisen, 2007; Bornmann; Daniel, 2008; Silveira; Caregnato, 2017; Aksnes; Langfeldt; Wouters, 2019; Tahamtan; Bornmann, 2022) a partir de, pelo menos, duas grandes vertentes: a normativa e a socioconstrutivista (ou, como nomeia Davis (2009), os campos da “citação como recompensa” e da “citação como persuasão”).

A teoria normativa, baseada na Sociologia da Ciência de Merton (1910-2003), enxerga a ciência como uma instituição social regida por normas e valores que orientam o comportamento dos pesquisadores. Nesse sentido, defende que os cientistas trocam informações em busca de reconhecimento, que se consolida na forma de prêmios e citações (Merton, 2013).

A visão sugere que as citações são uma forma de reconhecer dívidas intelectuais e, portanto, são influenciadas principalmente pelo valor percebido do trabalho, bem como pelo seu conteúdo cognitivo, metodológico ou temático. Ou seja, os cientistas citam porque isso faz parte da prática e do *ethos* científico e se baseiam na estrutura de recompensas da ciência. A decisão de citar, segundo essa visão,

SUMÁRIO

baseia-se no mérito do trabalho, sem levar em consideração aspectos pessoais do autor, como prestígio, afiliação institucional, religião ou identidade. A crítica a essa teoria argumenta que nem todas as influências são citadas e que, frequentemente, as citações ocorrem por razões particulares, por exemplo, autocitações e citações de caráter crítico (Bornmann; Daniel, 2008; Davis, 2009; Aksnes; Langfeldt; Wouters, 2019; Tahamtan; Bornmann, 2022).

A perspectiva socioconstrutivista, fundamentada nas abordagens dos estudos sociais da ciência, está associada a autores como Barry Barnes, David Bloor, Michel Callon, Karin Knorr Cetina e Bruno Latour. Esses pensadores enfocam o papel das comunidades científicas na construção do conhecimento e não consideram a citação como reconhecimento intelectual, mas como uma ferramenta estratégica de persuasão. Ou seja, ao citar determinado trabalho, o autor associa-se simbolicamente a ele; assim, desacreditar seu texto implica, indiretamente, desacreditar também as fontes que o sustentam (Bornmann; Daniel, 2008; Davis, 2009; Aksnes; Langfeldt; Wouters, 2019).

Nesse sentido, as citações são compreendidas como dispositivos retóricos utilizados para fortalecer argumentos e legitimar posições dentro do campo científico, funcionando de modo semelhante à retórica empregada em contextos de marketing ou publicidade (Davis, 2009). Ou seja, o propósito da citação não é reconhecer trabalhos anteriores, mas alcançar objetivos científicos, políticos e pessoais (Tahamtan, Bornmann, 2022).

Estudos empíricos não são conclusivos em relação ao papel de cada uma dessas teorias, portanto, o debate sobre o que as citações representam e, conseqüentemente, o que os estudos de citação medem continua. Nicolaisen (2007) sugere uma explicação sociobiológica, em que transparecer desonestidade afeta negativamente o capital científico do pesquisador, efeito que estimula que os autores façam citações adequadas (Tahamtan; Bornmann, 2022),

SUMÁRIO

por outro lado, integram trabalhos anteriores e, baseados em Niklas Luhmann, propõem uma Teoria das Citações como Sistemas Sociais (*Social Systems Citation Theory*).

Diante das distintas abordagens que buscam explicar o fenômeno da citação, observa-se que as citações podem refletir tanto o mérito cognitivo dos trabalhos quanto estratégias discursivas inseridas em contextos sociais e institucionais. Essas interpretações múltiplas evidenciam a complexidade do uso das citações como indicadores científicos.

Para que os estudos de citação esclareçam essa complexidade, é fundamental o acesso a dados estruturados e confiáveis sobre a produção científica. Nesse sentido, as fontes de informação desempenham um papel central para a coleta, a extração e o tratamento das citações em estudos métricos, conforme discutiremos nas seções a seguir.

FONTES DE INFORMAÇÃO PARA O ESTUDO DE CITAÇÃO: BASES DE DADOS BIBLIOGRÁFICAS E BUSCADORES DA LITERATURA CIENTÍFICA

As fontes de dados de citações são responsáveis por rastrear, compilar e contabilizar as citações recebidas pelos trabalhos que indexam. Esses serviços derivam das bibliografias especializadas, cujo desenvolvimento se expandiu significativamente na segunda metade do século XX, em resposta ao aumento da produção científica e à necessidade de acesso organizado à informação científica e tecnológica.

SUMÁRIO

Produzidos por associações profissionais, sociedades eruditas, agências governamentais, organizações internacionais e editoras comerciais, essas plataformas tornaram-se empreendimentos essenciais para a comunicação científica. De modo genérico, as fontes de informação relacionadas à literatura científica periódica podem ser classificadas da seguinte forma:

- a. Índices: resultados da indexação de artigos de periódicos, compreendem listagens ordenadas dos diferentes campos de metadados dos artigos científicos, como autor, título, periódico e assunto, servindo como instrumentos para a recuperação da literatura científica em diferentes áreas do conhecimento.
- b. Resumos: ampliam os índices ao incluir o resumo dos artigos como campo de indexação e busca, o que possibilita a recuperação por palavras da linguagem natural.
- c. Índices de citação: além dos metadados tradicionais, registram os itens presentes nas listas de referências, conectando obras citantes e citadas. Exemplos clássicos incluem o *Science Citation Index*, o *Social Sciences Citation Index* e o *Arts and Humanities Citation Index*, integrados à *Web of Science*, além da base *Scopus*.

Martín-Martín *et al.* (2021) citam o *Google Scholar*, *Web of Science*, *Scopus*, *Microsoft Academic*, *Dimensions* e *OpenCitations* como algumas das fontes de dados de citação mais relevantes e amplamente utilizadas. O *Google Scholar* e o *Microsoft Academic Graph* (descontinuado em 2022), apesar de serem mecanismos de busca e não bases de dados bibliográficas tradicionais, destacam-se por oferecer uma das coberturas mais amplas tanto de itens bibliográficos quanto de citações.

O ***Google Scholar*** é um mecanismo de busca voltado para conteúdos acadêmicos, indexando publicações de diversas naturezas

SUMÁRIO

e fontes, como artigos de periódicos, livros, teses, anais de eventos científicos, documentos institucionais e até material hospedado em páginas pessoais e repositórios. Seu algoritmo proprietário rastreia a web acadêmica de forma ampla, sem depender da curadoria formal ou de critérios editoriais rígidos como nas bases.

Entre suas principais vantagens do *Google Scholar* estão o acesso gratuito, a abrangência multidisciplinar e a indexação de documentos que muitas vezes não estão presentes em bases comerciais. No entanto, ele apresenta limitações: não possui transparência quanto aos critérios de indexação e atualização, oferece poucas opções de filtragem e exportação de dados, não permite buscas refinadas por campo bibliográfico, e os dados de citação podem conter duplicações, erros e inconsistências. Por esses motivos, embora útil como ferramenta de descoberta, o *Google Scholar* deve ser utilizado com cautela em estudos bibliométricos que exigem precisão, reprodutibilidade e controle da base analisada (Martín-Martín *et al.*, 2021).

A *Dimensions*, criada pela *Digital Science* em colaboração com diversas instituições, integra dados sobre publicações, citações, altmetria, ensaios clínicos, patentes e documentos de políticas públicas. Utiliza inteligência artificial para análise de tendências e impacto da pesquisa em escala global, com base no *Digital Object Identifier* (DOI) dos documentos. Possui filtros avançados de pesquisa, gráficos de citações e acesso a conteúdos em acesso aberto, apoiando o mapeamento do ciclo de vida da pesquisa (Universidade Estadual de Campinas, 2022; Universidade de São Paulo, [2022]).

O movimento da Ciência Aberta estimulou a criação de iniciativas que oferecem **citações abertas**, como a *OpenCitations*. A *CrossRef* passou a disponibilizar seus metadados de citação para coleta por terceiros, viabilizando projetos como o *OpenAlex* e os *OpenCitations Indexes*.

SUMÁRIO

OpenCitations Indexes são ferramentas desenvolvidas e disponibilizadas pela *OpenCitations*, que é uma infraestrutura para acesso a dados bibliográficos e de citações segundo os preceitos da Ciência Aberta (Heibi; Peroni; Shotton, 2019; Peroni; Shotton, 2019). Os diversos *OpenCitations Indexes* contêm metadados que ligam os documentos citados por meio de identificadores persistentes, mas não incluem metadados bibliográficos referentes às entidades citantes e citadas (como título, autores, número de páginas etc.). Para suprir essa deficiência, que dificultava levantamentos de citações e estudos bibliométricos, a organização lançou recentemente o **OpenCitations Meta**, uma database que complementa os índices, oferecendo metadados bibliográficos das publicações científicas registradas no *Crossref*, *DataCite*, e *PubMed*, também de forma aberta (Massari *et al.*, 2024).

O **OpenAlex**, desenvolvido pela organização sem fins lucrativos *OurResearch*, é uma base bibliográfica de código aberto (licença *Creative Commons CC0*) cujos dados originais foram incorporados do legado do *Microsoft Academic Graph*. Ele é atualizado constantemente com dados do *Crossref*, *Pubmed*, *Open Researcher and Contributor ID* (ORCID), *Research Organization Registry* (ROR), *Directory of Open Access Journals* (DOAJ), *arXiv*, *Zenodo*, entre outros (Massari *et al.*, 2024). Indexa publicações, autores, instituições e dados de financiamento, fornecendo indicadores de citação e visualizações gráficas das relações entre esses elementos. Estimula o compartilhamento e reuso de dados, alinhando-se aos princípios da Ciência Aberta (Universidade de São Paulo, c2025).

O uso de identificadores únicos e permanentes, como o DOI e o *Handle System*, assegura a integridade dos recursos informacionais, mesmo em ambientes digitais distribuídos. Assim como o *International Standard Book Number* (ISBN) identifica livros e o *International Standard Serial Number* (ISSN) identifica publicações seriadas de forma inequívoca, esses sistemas identificam objetos digitais, estabelecendo conexões entre metadados e conteúdo.

SUMÁRIO

A **CrossRef**, associação independente de editores científicos, se destaca nesse cenário ao fornecer infraestrutura para a atribuição de DOIs, além de promover a interoperabilidade por meio de links referenciais. O DOI permite a integração entre publicações acadêmicas eletrônicas, pois faz com que os editores depositem, com as URLs, também os metadados dos artigos, como título, autor, periódico, ano de publicação, além das referências dos documentos citados, o que permite criar bases de dados abertas de citação, como o *Open Citation Index* e o *OpenAlex*.

O **Handle System** sustenta o próprio DOI e outros sistemas de identificação. Embora tenha viés comercial, o DOI pode ser usado para qualquer tipo de propriedade intelectual digital, incluindo conteúdos abertos. Sua funcionalidade clicável estabelece elos entre recursos na web, facilitando sua localização e reutilização. No quadro 1, são apresentadas as principais bases de dados que permitem a extração de citações e metadados relacionados por meio de *Application Programming Interfaces* (API).

Quadro 1 – Principais bases de dados para extração de citações via API

Nome da base	Tipo acesso à API
<i>Web of Science</i>	Pago
<i>Scopus</i>	Pago
<i>OpenAlex</i>	Gratuito
<i>Dimensions</i>	Pago (com possibilidade de gratuidade mediante análise)
<i>Lens.org</i>	Pago
<i>OpenCitations</i>	Gratuito
<i>Crossref</i>	Gratuito*

Fonte: elaboração própria.

Nota: *uma versão avançada requer assinatura.

SUMÁRIO

A escolha e o uso das fontes de informação para estudos de citação são determinantes para a qualidade e a confiabilidade das análises realizadas. Isso porque a utilidade dessas fontes depende diretamente da qualidade dos metadados de citação que elas oferecem. A próxima seção aborda a qualidade de metadados de citação com mais detalhes.

QUALIDADE DOS METADADOS DE CITAÇÃO

Para que os estudos de citação correspondam às práticas dos pesquisadores, é necessário que os metadados que descrevem os documentos apresentem qualidade no que se refere ao conjunto de características que devem ser observadas em qualquer domínio. Segundo Bruce e Hillmann (2004), essas características incluem: completude, precisão, proveniência, conformidade com as expectativas, consistência e coerência lógica, atualização e acessibilidade. O quadro 2 mostra um resumo das características descritas por Bruce e Hillmann (2004) relacionadas à qualidade dos metadados.

Quadro 2 – Características de qualidade dos metadados, segundo Bruce e Hillmann (2004)

Critério	Descrição
Completude	Os metadados devem descrever os objetos de forma completa, incluindo informações essenciais para identificação, localização e compreensão, mantendo, porém, o equilíbrio entre quantidade de informações e viabilidade de sua implementação. Adicionalmente, o esquema de metadados deve ser aplicável a toda a coleção de objetos.
Precisão	Essa característica estabelece que os metadados devem ser corretos e factuais, sendo o mais exatos possível, devendo-se evitar erros tipográficos, uso inconsistente de abreviações ou discrepâncias na escrita de nomes.
Proveniência	A proveniência de metadados diz respeito à sua origem, quem preparou, como foram criados e quaisquer transformações ou tratamentos aplicados.

SUMÁRIO

Critério	Descrição
Conformidade com as expectativas	Os metadados devem conter os elementos que a comunidade esperaria encontrar. Isso significa que os padrões de metadados devem atender às demandas práticas do público-alvo, evitando elementos irrelevantes.
Consistência e coerência lógica	Os elementos devem ser concebidos de forma consistente com definições e conceitos padrão, principalmente em coleções heterogêneas. É enfatizado o uso de mecanismos padronizados para representação dos metadados.
Atualização	A atualização é abordada sob dois aspectos: a capacidade dos metadados de se manterem atualizados conforme as mudanças ocorridas no objeto; e o atraso, que diz respeito ao tempo decorrido entre a disseminação (ou criação/disponibilização) do objeto e o momento em que os metadados correspondentes são gerados ou disponibilizados. Este último se aplica, mas não se limita, aos casos de registros de citações, em que há um atraso natural entre a publicação do objeto citante e sua contabilização nos metadados do objeto citado. Quanto menor o atraso, maior a confiança e a qualidade dos metadados.
Acessibilidade	A acessibilidade dos metadados significa que eles devem ser facilmente acessíveis, legíveis e compreendidos pelos usuários. Os obstáculos que podem prejudicar a acessibilidade incluem: problemas técnicos (por exemplo, metadados separados dos objetos que descrevem ou armazenados em formatos obsoletos), barreiras intelectuais (como uso de vocabulários ou sistemas de classificação que fazem sentido apenas para especialistas) e obstáculos econômicos (o acesso só é possível mediante assinaturas ou custos adicionais).

Fonte: elaboração própria, com base em Bruce e Hillmann (2004).

Em análises de citação, os metadados necessários e frequentemente requisitados incluem o número de citações e a relação dos trabalhos citantes. A qualidade e a confiabilidade desses metadados exige, no mínimo, a observância dos seguintes critérios: precisão, consistência, atualização e acessibilidade.

Quanto à precisão e consistência, é importante que exista correspondência entre a quantidade de citações e quantidade de citantes listados pelo sistema ou base de dados, bem como a presença do citado nas referências bibliográficas dos documentos citantes. Quanto ao critério de atualização é recomendado que a base de dados possua mecanismos de identificação e incorporação de metadados em intervalos regulares, preferencialmente de curto prazo.

SUMÁRIO

A acessibilidade desses metadados é importante para estudos de citação, devendo, portanto, estarem disponíveis em um formato legível tanto por computadores quanto por humanos. Além disso, garantir a persistência e a disponibilidade desses dados a longo prazo, possibilitando verificações futuras é fundamental. Adicionalmente, considera-se importante que as bases de dados incluam registros históricos das citações, estratificados por períodos (anuais, semestrais, mensais...), permitindo análises temporais das citações.

O fornecimento e atualização dos metadados de citação dependem em grande parte dos periódicos ou editoras responsáveis pela publicação dos documentos, para os quais devem disponibilizar metadados ricos. Metadados ricos são a descrição completa e detalhada de um objeto, garantindo que suas informações sejam armazenadas em uma base de dados que permita sua recuperação, interoperabilidade e durabilidade, sem comprometer as qualidades destacadas por Bruce e Hillmann (2004). O enriquecimento de metadados, segundo di Buono *et al.* (2022), visa melhorar a informação dos recursos e pode ser feito de forma automática ou manual, com o objetivo de encontrar e corrigir inconsistências e valores ausentes nos metadados.

Nesse contexto, é importante que os editores de periódicos priorizem metadados ricos em nível de artigo. Eles devem incluir o máximo de elementos descritivos possível aos metadados para atender às necessidades de diferentes interessados, além de garantir que os metadados sejam legíveis por máquina, mantendo-os organizados, consistentes e interoperáveis. Para isso, as revistas podem coletar metadados através dos formulários de submissão de artigos e gerar metadados *Extensible Markup Language* (XML) em nível de artigo na linguagem de marcação *Journal Article Tag Suite* (JATS) (In Pursuit..., 2021).

SUMÁRIO

Como destaca o texto *In Pursuit...* (2021), além dos metadados básicos (como título do periódico, título do artigo, autoria etc.) há, atualmente, cinco elementos essenciais que editores devem priorizar:

- a. **Identificadores Persistentes (PIDs):** embora o DOI seja amplamente utilizado na comunicação científica, existem outros identificadores persistentes igualmente relevantes e que permitem a identificação inequívoca de diversas entidades. Entre eles, destaca-se o ROR, utilizado para identificar instituições; o ORCID, voltado para autores e contribuintes; e os identificadores de bolsas e financiadores, como o *Open Funder Registry* (OFR) (*In Pursuit...*, 2021).
- b. **Licenças de direitos autorais:** informações sobre licenças de direitos autorais são importantes para que os periódicos atendam às iniciativas de *Open Access*. Além disso, podem contribuir para aumentar a visibilidade dos artigos, especialmente em periódicos de acesso aberto, uma vez que diversos indexadores e serviços de busca possibilitam a pesquisa e filtragem com base na licença de direitos autorais (*In Pursuit...*, 2021).
- c. **Resumos de artigos:** incluir os resumos dos artigos nos metadados fornece às bases de dados e serviços de busca informações descritivas detalhadas, facilitando o “entendimento” do conteúdo por essas ferramentas. Além disso, esses resumos legíveis por máquina, abrem diversas possibilidades para a mineração de textos e dados (*In Pursuit...*, 2021). De acordo com a *Initiative for Open Abstracts* (I4OA), disponibilizar os resumos abertamente permite ampliar a visibilidade e o impacto dos periódicos e livros. Ademais, facilita a descoberta, leitura e citação dessas publicações, contribui para sua inclusão em revisões sistemáticas e permite o uso de técnicas como mineração de texto, processamento de linguagem natural e inteligência artificial em análises bibliométricas (*Initiative for Open Abstracts...*, 2025).

SUMÁRIO

- d. **Citações abertas:** incluir citações abertas significa que os editores devem depositar a lista de referências como parte dos metadados do artigo na *Crossref* e permitir sua distribuição, tornando-as abertamente disponíveis por padrão (In Pursuit..., 2021). Assim, a lista de referências dos artigos torna-se acessível para qualquer interessado por meio dos serviços de entrega de metadados da *Crossref*, como a *API REST*, e passa a integrar o *Cited-by*, serviço que informa a contagem e a lista de citações de um determinado trabalho (Tolwinska, 2018). Dessa forma, depositar as listas de referências na *Crossref* é importante, pois permite sua leitura por máquina, facilitando estudos de citações e possibilitando a identificação e o rastreamento de conexões entre as pesquisas.
- e. **Principais meta tags HyperText Markup Language (HTML) para o Google Scholar:** *meta tags* são elementos HTML que fornecem informações sobre uma página da web para os mecanismos de busca e navegadores (Google Search Central, [2025]). As meta tags não são visíveis na página, mas são lidas por robôs. O *Google Scholar* utiliza essas *tags* para identificar e organizar os artigos. Para que um artigo seja indexado pelo *Google Scholar*, ele deve conter, no mínimo, três *meta tags* HTML, cada uma com as seguintes informações: título do artigo; nome completo de pelo menos um dos autores e ano de publicação. Além disso, cada artigo precisa estar em uma página web própria, mesmo que seja um arquivo Portable Document Format (PDF). Além dos três elementos obrigatórios, outras *meta tags* podem ser utilizadas para enriquecer os metadados, como: nome da revista; DOI e a *Uniform Resource Locator* (URL) do artigo (In Pursuit..., 2021).

Em resumo, a qualidade e a riqueza dos metadados são importantes não apenas para recuperação da informação, mas também para estudos de citação. Garantir a qualidade e a riqueza dos metadados de publicação é uma responsabilidade coletiva que recai,

SUMÁRIO

principalmente, sobre os autores, para que forneçam informações fidedignas sobre o trabalho; sobre os editores, quanto ao fornecimento, preenchimento, depósito, manutenção e consistência dos metadados; e sobre as bases de dados e ferramentas de busca que ficam responsáveis pela guarda, edição, fornecimento de acesso, interoperabilidade, bem como pela persistência e preservação dos dados.

FERRAMENTAS PARA EXTRAÇÃO DE CITAÇÕES

Extraír grandes volumes de citações demanda o uso de ferramentas automatizadas, cuja escolha dependerá do formato dos dados e objetivos da pesquisa. Existem diversos métodos e procedimentos para a extração de dados de citação, sendo o uso de API a forma comumente oferecida pelas bases de dados.

Segundo Pawar *et al.* (2023), uma API é um conjunto de regras, métodos e ferramentas utilizadas para criar funcionalidades em sistemas de software. Em outras palavras, a API serve como uma ponte entre diferentes programas, permitindo que eles se comuniquem. Assim, ao contrário das interfaces de usuário, as API permitem que softwares ou computadores se comuniquem entre si, trocando dados e informações. Ela torna visíveis os métodos (ou comandos), os dados de entrada necessários, os resultados e os tipos de dados que fazem parte de um software. Dessa forma, ao utilizar uma API, é possível acessar certas funcionalidades de um sistema sem precisar entender ou modificar sua estrutura interna (Pawar *et al.*, 2023).

Nesta seção, focou-se em como extrair dados da base *OpenAlex*, pois esta é uma fonte completamente aberta de metadados acadêmicos, que possui potencial de aprimorar a transparência na avaliação, navegação, representação e descoberta científica.

SUMÁRIO

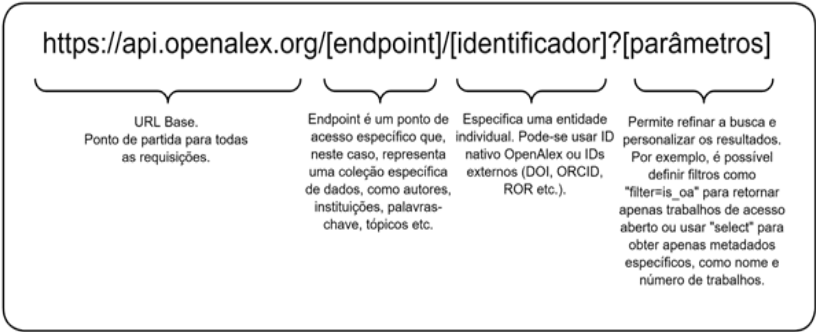
Atualmente, o *OpenAlex* se figura como uma das maiores bases de dados abertos. A plataforma, que incorporou os dados do *Microsoft Academic Graph* (MAG), reúne mais de 209 milhões de trabalhos, cerca de 213 milhões de autores, mais de 124 mil veículos entre periódicos, conferências, repositórios de *preprints* e repositórios institucionais e aproximadamente 109 mil instituições (Priem; Piwowar; Orr, 2022).

Além disso, apesar de seu recente lançamento (em 2021), o *OpenAlex* vem sendo utilizado cada vez mais como ferramenta para análises bibliométricas. Scheidsteger e Haunschild (2023), por exemplo, compararam a qualidade dos metadados do *OpenAlex* com a dos metadados do MAG e concluíram que o *OpenAlex* se mostrou tão adequado para análises bibliométricas quanto seu antecessor. Nesse mesmo sentido, Alperin *et al.* (2024) compararam a cobertura de dados do *OpenAlex* com a da *Web of Science* e constataram que o *OpenAlex* oferece uma cobertura mais ampla e inclusiva, constituindo-se, assim, em uma fonte alternativa para muitas análises.

Segundo o próprio site, o *OpenAlex* tem como objetivo democratizar o conhecimento científico por meio da Ciência Aberta, oferecendo uma alternativa às bases de dados pagas, como a *Scopus* e a *Web of Science* (OpenAlex Support, [2023-]). Sua coleção é composta por trabalhos acadêmicos, incluindo artigos, livros, conjuntos de dados e dissertações, cujos metadados são disponibilizados gratuitamente sob licença *Creative Commons* CC0 através de sua API. Os dados são coletados de diversas fontes, incluindo *Crossref*, *DataCite*, *PubMed*, *DOAJ*, *ORCID*, *Microsoft Academic Graph* (descontinuado), *ISSN*, repositórios institucionais, e de milhares de outras fontes (OpenAlex Support, [2023-]).

O acesso à API é realizado por meio de requisições HTTP GET, método usado para solicitar informações de um servidor na internet. Dessa forma, a interação com a API ocorre por meio da construção de URLs específicas, conforme os dados que se deseja obter. A construção das URLs da API *OpenAlex* segue uma estrutura predefinida, conforme mostrado na figura 1:

Figura 1 - Estrutura básica das URLs da API OpenAlex



Fonte: elaboração própria, com base em OpenAlex Technical Documentation (2025).

Na estrutura da URL, o *endpoint*, o identificador e os parâmetros são variáveis, o que possibilita inúmeras formas de recuperar os dados. O quadro 3 detalha cada uma dessas variáveis e apresenta exemplos de aplicação.

Quadro 3 – Estrutura de URLs na API OpenAlex: endpoints, identificadores e parâmetros

Endpoint: especifica o tipo de entidade que deseja acessar		
Principais Variáveis	Descrição	Exemplo
/works	Retorna todos os trabalhos acadêmicos (artigos, livros, dissertações etc.) disponíveis na base.	https://api.openalex.org/works
/authors	Retorna todos os autores disponíveis.	https://api.openalex.org/authors
/sources	Retorna todas as fontes (revistas, repositórios, conferências etc.) disponíveis.	https://api.openalex.org/sources
/institutions	Retorna todas as instituições disponíveis.	https://api.openalex.org/institutions
/topics	Retorna tópicos (assuntos) relacionados aos trabalhos disponíveis.	https://api.openalex.org/topics
/publishers	Refere-se a editoras.	https://api.openalex.org/publishers
/funders	Retorna agências financiadoras.	https://api.openalex.org/funders

SUMÁRIO

Identificador ou lista de identificadores: para obter informações sobre uma entidade específica, geralmente adiciona-se o identificador dessa entidade à URL		
Identificadores aceitos	Descrição	Exemplo
<i>OpenAlex</i> IDs	Identificadores únicos atribuídos pelo <i>OpenAlex</i> no momento da indexação do item. Eles começam com uma letra que indica o tipo da entidade, por exemplo: W para <i>works</i> , A para <i>authors</i> , S para <i>sources</i> , etc.	https://api.openalex.org/works/W2741809807 https://api.openalex.org/authors/A5048491430
IDs externos	Como DOI (para trabalhos), ORCID (para autores), ISSN (para fontes), ROR (para instituições), <i>Wikidata</i> ID (para conceitos e tópicos), PMID (<i>PubMed Unique Identifier</i>), etc. IDs externos devem ser apresentados com a URL completa.	https://api.openalex.org/works/https://doi.org/10.7717/peerj.4375 https://api.openalex.org/authors/https://orcid.org/0000-0003-1613-5981 https://api.openalex.org/institutions/https://ror.org/0130frc33
Principais parâmetros		
Parâmetro	Descrição	Exemplo
filter	Usado para filtrar os resultados com base em atributos específicos. Por exemplo, para obter apenas trabalhos de acesso aberto. É possível usar múltiplos filtros separados por vírgula (,).	https://api.openalex.org/works?filter=is_oa:true
select	Usado para selecionar quais campos ou metadados serão incluídos na resposta, reduzindo o tamanho da resposta. Os campos são separados por vírgula.	https://api.openalex.org/works/W2741809807?select=id,title
per_page	Usado para especificar o número de resultados por página. É possível definir o parâmetro <i>per_page</i> para qualquer valor entre 1 e 200.	https://api.openalex.org/works?per_page=200

SUMÁRIO

page	Usado para navegar entre as diferentes páginas de resultados.	https://api.openalex.org/works?page=2
search	Utilizado para realizar uma busca textual dentro do <i>endpoint</i> de um específico da API do <i>OpenAlex</i> .	https://api.openalex.org/publishers?search=springer
group_by	Este parâmetro permite agrupar e contar entidades que compartilham o mesmo valor para um determinado atributo. Pode ser utilizado para obter agregações de dados.	https://api.openalex.org/works?group_by=open_access.oa_status https://api.openalex.org/institutions?group_by=country_code

Fonte: elaboração própria, com base em *OpenAlex Technical Documentation* (2025).

O quadro 3 apresenta uma diversidade de *endpoints*, permitindo que os usuários recuperem informações conforme os mais variados objetivos de pesquisa. A combinação de *endpoints*, identificadores e parâmetros possibilita direcionar a extração e recuperar somente os dados relevantes para a pesquisa. A título de exemplo: é possível combinar o *endpoint* “works” com o parâmetro “filter” para obter todos os trabalhos de um determinado ano (a URL, nesse caso, ficaria assim: https://api.openalex.org/works?filter=publication_year:2020). Também é possível combinar o *endpoint* *topics* com o parâmetro *select*, especificando os metadados desejados - como o ID *OpenAlex*, a denominação do tópico/assunto, a quantidade de trabalhos publicados e o número de citações recebidas. A URL para esse exemplo seria: https://api.openalex.org/topics?select=id,display_name,works_count,cited_by_count.

Observa-se o emprego de caracteres como “?”, “=”, “/”, “&”, “:” e “,”. Esses caracteres fazem parte da sintaxe das URLs e desempenham funções importantes, sendo as principais:

- a. Ponto de interrogação (?): tem a função de ser um delimitador da consulta. Tudo o que vem após o “?” são parâmetros enviados à API, com o objetivo de refinar ou delimitar

SUMÁRIO

os dados que se deseja obter. Deve ser utilizado sempre que se quiser adicionar um parâmetro à URL, e deve ser colocado antes dele. Exemplo: <https://api.openalex.org/works/W2741809807?select=id,title>. Nesse exemplo o “?” delimita que está sendo inserido o parâmetro *select* na *string* (refere-se ao endereço URL da API estruturado com os parâmetros, filtros e *endpoints* desejados).

- b. Sinal de igual (=): serve para atribuir um determinado valor a um parâmetro específico. Exemplo: <https://api.openalex.org/works/W2741809807?select=id,title> - nesse caso, esse sinal indica que o parâmetro *select* deve retornar apenas os campos/valores *id* e *title* na API.
- c. Sinal de barra (/): serve para separar os segmentos da URL, sendo usado principalmente antes dos *endpoints* e antes de identificadores, como por exemplo: <https://api.openalex.org/works/W2741809807>.
- d. Ampersand (&): é um separador de parâmetros e deve ser utilizado sempre que se deseja incluir mais de um parâmetro na URL. Exemplo: na URL https://api.openalex.org/works?per_page=200, o parâmetro *per_page* define o número de resultados exibidos por página. Caso se deseje acessar a segunda página de resultados, é necessário acrescentar o parâmetro *page*, seguido do sinal de igual e do número da página desejada, ou seja, *page=2*. Para combinar ambos os parâmetros, utiliza-se o “&” entre eles, resultando na seguinte URL: https://api.openalex.org/works?per_page=200&page=2.
- e. Dois pontos (:): funciona como um indicador do valor que se deseja do campo ao qual é precedido. Por exemplo na URL https://api.openalex.org/works?filter=publication_year:2020. O sinal de dois pontos está indicando qual o ano deve ser considerado para o campo “*publication_year*”. Para buscar um intervalo de anos

SUMÁRIO

consecutivos pode-se usar o caractere “-”, por exemplo https://api.openalex.org/works?filter=publication_year:2020-2024.

- f. Vírgula (,): na sintaxe da URL, a vírgula funciona como um delimitador de lista e é usada principalmente no parâmetro *select*, quando se deseja selecionar vários campos diferentes. Exemplo: <https://api.openalex.org/works/W2741809807?select=id,title>, onde a vírgula separa os campos *id* e *title* que devem ser retornados.

A vírgula também pode ser utilizada para combinar múltiplos filtros dentro de uma mesma *string*. Exemplo: https://api.openalex.org/works?filter=publication_year:2020-2024,institutions.id:https://openalex.org/i1294671590. Nesse caso, a busca aplica dois filtros simultaneamente - um para o intervalo de anos de publicação e outro para o ID de uma instituição. Assim, nesse exemplo, a vírgula atuou como um operador lógico “AND”, ou seja, a API retorna apenas os resultados que atendem a ambas as condições.

Esses são apenas alguns dos caracteres que podem ser utilizados na estruturação da *string* da API *OpenAlex*. O uso deles é variado e não se restringe apenas aos exemplos apresentados acima.

Para a coleta de dados do *OpenAlex* por meio da sua API, é fundamental, além da construção adequada da *string*, conhecer também os metadados disponíveis. Os principais metadados relacionados às citações que se pode encontrar no *OpenAlex* podem ser classificadas em três categorias:

- a. **Metadados de contagem de citações:** inclui dois metadados, o *cited_by_count*, que indica o número de citações recebidas, e o *counts_by_year*, que mostra a quantidade de citações por ano.
- b. **Metadados relacionados a métricas de impacto:** são apresentados cinco metadados vinculados a métricas, que

SUMÁRIO

incluem: o *Field-Weighted Citation Impact* (FWCI), que mede a proporção entre as citações recebidas e as citações esperadas, com base na média de citações de cada campo de estudo; o *citation_normalized_percentile*, que apresenta as mesmas informações do FWCI, mas expressas em forma de percentil; o *2yr_mean_citedness*, que representa a média de citações dos últimos 2 anos, também conhecido como Fator de Impacto; o *h_index*, que é o índice H do autor, fonte, instituição etc.; e o *i10_index*, que corresponde ao número de publicações com pelo menos 10 citações.

- c. **Metadados de referências e citantes:** também é possível acessar a lista de referências e a lista de citantes. Isso se dá por meio dos metadados: *referenced_works*, que apresenta os trabalhos referenciados por um determinado trabalho; e *cited_by_api_url*, que indica uma URL da API onde se encontra a listagem de todos os citantes daquele trabalho, juntamente com os metadados de cada citante.

Quanto à extração desses metadados, é possível realizá-la de forma automática por meio de requisições HTTP em diversas linguagens de programação, como *Python*, *R*, *JavaScript* e *Java*. Esse método permite a automação do processo, possibilitando a obtenção de grandes volumes de dados, que, após extraídos, podem ser salvos em formatos de planilha como CSV e XLS ou ainda em formato de banco de dados como *SQLite* ou *SQL*.

Uma forma eficaz e prática de extrair dados do *OpenAlex* é por meio do software *RStudio*, um ambiente de desenvolvimento integrado (IDE) projetado para facilitar o uso da linguagem de programação *R*. O *RStudio* é voltado especialmente para apoiar cientistas de dados nas mais diversas atividades (Posit Software, 2025). Ele conta com um pacote denominado *OpenalexR*, que reúne um conjunto de ferramentas para extrair dados bibliográficos do

SUMÁRIO

OpenAlex por meio de sua API (*OpenalexR*, 2025). Com esse pacote, é possível obter os metadados das seguintes entidades: trabalhos ("*works*"), autores ("*authors*"), instituições ("*institutions*"), palavras-chave ("*keywords*"), financiadores ("*funders*"), fontes ("*sources*"), publicadores ("*publishers*") e tópicos ("*topics*").

A disponibilização aberta desses metadados permite a realização de diversos estudos, como análises de impacto acadêmico, identificação de autores e artigos mais influentes em cada campo, a evolução das citações ao longo do tempo e redes de citações que possibilitam identificar grupos de pesquisadores e instituições que atuam nos mesmos temas ou que frequentemente são citados em conjunto. Ademais, a disponibilidade desses dados viabiliza estudos comparativos que avaliam as variações nos padrões de citação em diferentes países, áreas de conhecimento e idiomas.

CONSIDERAÇÕES FINAIS

As citações constituem um dos elementos principais da comunicação científica, pois articulam a construção coletiva do conhecimento, conectando autores, textos, ideias e instituições ao longo do tempo. Embora sejam amplamente utilizadas como indicadores de impacto, sua natureza é multifacetada, combinando dimensões conceituais, sociais, institucionais e retóricas. Por isso, os estudos de citação necessitam de fundamentação teórica consistente, compreensão crítica das motivações que levam os autores a citar e uso criterioso das métricas derivadas.

Este capítulo buscou apresentar os principais fundamentos conceituais e metodológicos para os estudos de citação, destacando as fontes de dados disponíveis, com ênfase nas bases abertas, e as ferramentas que viabilizam a extração e análise de citações. Discutiu-se também a importância da qualidade dos metadados,

SUMÁRIO

que devem ser completos, precisos, atualizados e acessíveis para garantir a confiabilidade das análises.

O movimento da Ciência Aberta ampliou as possibilidades de realização de estudos bibliométricos mais transparentes, reprodutíveis e inclusivos. Iniciativas como *OpenAlex* e *OpenCitations* oferecem alternativas promissoras às bases comerciais, democratizando o acesso a dados de citação e promovendo a abertura dos fluxos informacionais que sustentam a avaliação da atividade científica.

Contudo, para que os estudos de citação alcancem seu pleno potencial, é necessário um esforço conjunto de autores, editores, bibliotecários e instituições de pesquisa no sentido de enriquecer, padronizar e disponibilizar metadados com qualidade e de forma interoperável. Assim, será possível não apenas recuperar informações com maior precisão, mas, também, promover uma compreensão mais ampla e crítica dos processos de produção e circulação do conhecimento científico.

REFERÊNCIAS

OpenAlex Support. **About us**. [S. /], [2023-]. Disponível em: <https://help.openalex.org/hc/en-us/articles/24396686889751-About-us>. Acesso em: 30 jul. 2025.

ALPERIN, J. P. *et al*. An analysis of the suitability of OpenAlex for bibliometric analyses. **arXiv.org**, [S. /], 2024. Preprint. Disponível em: <https://doi.org/10.48550/arXiv.2404.17663>. Acesso em: 30 jul. 2025.

AKSNES, D. W.; LANGFELDT, L.; WOUTERS, P. Citations, citation indicators, and research quality: An overview of basic concepts and theories. **SAGE Open**, Thousand Oaks, v. 9, n. 1, p. 1-17, Jan./Mar. 2019. Disponível em: <https://doi.org/10.1177/2158244019829575>. Acesso em: 30 jul. 2025.

BORNMANN, L.; DANIEL, H.-D. What do citation counts measure? A review of studies on citing behavior. **Journal of Documentation**, London, v. 64, n. 1, p. 45-79, 2008. Disponível em: <https://doi.org/10.1108/00220410810844150>. Acesso em: 30 jul. 2025.

SUMÁRIO

BRUCE, T. R.; HILLMANN, D. I. **The continuum of metadata quality**: defining, expressing, exploiting. Chicago: ALA Editions, 2004.

CRONIN, B. **The citation process: the role and significance of citations in scientific communication**. London: Taylor Graham, 1984.

DAVIS, P. M. Reward or persuasion? The battle to define the meaning of a citation. **Learned Publishing**, Watford, v. 22, n.1, p. 5-11, Jan. 2009. Disponível em: <https://doi.org/10.1087/095315108X378712>. Acesso em: 30 jul. 2025.

DI BUONO, M. P. *et al.* Paving the way for enriched metadata of linguistic linked data. **Semantic Web**, [S. /], v. 13, p. 1133-1157, 2022. Disponível em: <https://doi.org/10.3233/SW-222994>. Acesso em: 30 jul. 2025.

GARFIELD, E. Citation indexes for science: A new dimension in documentation through association of ideas. **Science**, Washington, D.C., v. 122, n. 3159, p. 108-111, July 1955. Disponível: <https://doi.org/10.1126/science.122.3159.108>. Acesso em: 30 jul. 2025.

GOOGLE SEARCH CENTRAL. **Meta tags e atributos compatíveis com o Google**. [S. /], [2025]. Disponível em: <https://developers.google.com/search/docs/crawling-indexing/special-tags?hl=pt-br>. Acesso em: 30 jul. 2025.

HEIBI, I., PERONI, S.; SHOTTON, D. Software review: COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. **Scientometrics**, Amsterdam, n. 121, p. 1213-1228, 2019. <https://doi.org/10.1007/s11192-019-03217-6>. Acesso em: 30 jul. 2025.

IN PURSUIT of rich article-level metadata: 5 elements journal publishers should prioritize. **Scholastica**, 8 June 2021. Disponível em: https://blog.scholasticahq.com/post/metadata-elements-publishers-should-prioritize/?utm_campaign=Organic+Social_Content&utm_source=twitter&utm_medium=social. Acesso em: 30 jul. 2025.

INITIATIVE FOR OPEN ABSTRACTS. [S. /], 2025. Disponível em: <https://i4oa.org>. Acesso em: 30 jul. 2025.

MARTÍN-MARTÍN, A. *et al.* Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. **Scientometrics**, Amsterdam, v. 126, p. 871-906, 2021. Disponível em: <https://doi.org/10.1007/s11192-020-03690-4>. Acesso em: 30 jul. 2025.

MASSARI, A. *et al.* OpenCitations Meta. **Quantitative Science Studies**, [S. /], v. 5, n. 1, p. 50-75, 2024. Disponível em: https://doi.org/10.1162/qss_a_00292. Acesso em: 30 jul. 2025.

SUMÁRIO

MERTON, R. K. **Ensaio sobre sociologia da ciência**. Tradução de Tradução: Sylvia Gemignani Garcia e Pablo Rubén Mariconda. São Paulo: Editora 34, 2013. (Filosofia da Ciência e da Tecnologia).

NICOLAISEN, J. Citation analysis. **Annual Review of Information Science and Technology**, Baltimore County, v. 41, n. 1, p. 609-641, 2007.

OPENALEX TECHNICAL DOCUMENTATION. **Overview**: OpenAlex. [S. /], 2025. Disponível em: <https://docs.openalex.org>. Acesso em: 30 jul. 2025.

OPENALEXR: getting bibliographic records from 'OpenAlex' database using 'DSL' API. [S. /], 4 May 2025. Versão 2.0.1. Disponível em: <https://cran.r-project.org/package=openalexR>. Acesso em: 30 jul. 2025.

PERONI, S.; SHOTTON, D. OpenCitations, an infrastructure organization for open scholarship. **Quantitative Science Studies**, [S. /], n. 1, v. 1, p. 428-444, 2020. Disponível em: https://doi.org/10.1162/qss_a_00023. Acesso em: 30 jul. 2025.

PRIEM, J.; PIWOWAR, H. A.; ORR, R. A. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. **arXiv.org**, [S. /], June 2022. Versão 2. Preprint. Disponível em: <https://doi.org/10.48550/arXiv.2205.01833>. Acesso em: 30 jul. 2025.

PAWAR, S. *et al.* Application programming interface with a case study of SOA. In: INTERNATIONAL CONFERENCE ON INTEGRATED INTELLIGENCE AND COMMUNICATION SYSTEMS, 24-25 Nov. 2023, Kalaburagi. **Proceedings...** 2023. p. 1-5. Disponível em: <https://doi.org/10.1109/ICIICS59993.2023.10421593>. Acesso em: 30 jul. 2025.

POSIT SOFTWARE. **RStudio desktop: download**. [S. /], 5 June 2025. Versão 2025.05.1+513. Disponível em: <https://posit.co/download/rstudio-desktop/>. Acesso em: 30 jul. 2025.

SCHEIDSTEGE, T.; HAUNSCHILD, R. Which of the metadata with relevance for bibliometrics are the same and which are different when switching from Microsoft Academic Graph to OpenAlex?. **Profesional de la Información**, London, v. 32, n. 2, 2023. Disponível em: <https://doi.org/10.3145/epi.2023.mar.09>. Acesso em: 30 jul. 2025.

SILVEIRA, M. A. A. da; CAREGNATO, S. E.; BUFREM, L. S. Estudo das razões das citações na Ciência da Informação: proposta de classificação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasília, DF, v. 7, n. 2, jul./dez. 2014. Disponível em: <https://ancib.org/revistas/index.php/tpbci/article/view/314>. Acesso em: 30 jul. 2025.

SUMÁRIO

SILVEIRA, M. A. A. da; CAREGNATO, S. E. Demarcações epistemológicas dos estudos de citação: o fenômeno da citação. **Informação & Sociedade: Estudos**, João Pessoa, v. 27, n. 3, set./dez. 2017. Disponível em: <https://doi.org/10.22478/ufpb.1809-4783.2017v27n3.33847>. Acesso em: 30 jul. 2025.

SMALL, H. On the shoulders of Robert Merton: Towards a normative theory of citation. **Scientometrics**, Amsterdam, v. 60, p. 71-79, 2004. Disponível em: <https://doi.org/10.1023/B:SCIE.0000027310.68393.bc>. Acesso em: 30 jul. 2025.

TAHAMTAN, I.; BORNMANN, L. The Social Systems Citation Theory (SSCT): A proposal to use the social systems theory for conceptualizing publications and their citation links. **Profesional de la Información**, Birmingham, v. 31, n. 4, e310411, 2022. Disponível em: <https://doi.org/10.3145/epi.2022.jul.11>. Acesso em: 30 jul. 2025.

TOLWINSKA, A. **Linking references is different from registering references**. [S. /], 30 May 2018. Disponível em: <https://doi.org/10.64000/ve3q2-3xm59>. Acesso em: 30 jul. 2025.

UNIVERSIDADE DE SÃO PAULO. Acesso aberto. **OpenAlex**: base bibliográfica analítica e aberta. São Paulo, c2025. Disponível em: <https://www.acessoaberto.usp.br/openalex-catalogo-e-analise-bibliometrica/>. Acesso em: 30 jul. 2025.

UNIVERSIDADE DE SÃO PAULO. Agência de bibliotecas e coleções digitais. **Plataforma Dimensions Analytics está disponível na Universidade de São Paulo (USP)**. São Paulo, [2022]. Disponível em: <https://www.abcd.usp.br/noticias/plataforma-dimensions-disponivel-na-universidade-de-sao-paulo-usp/>. Acesso em: 30 jul. 2025.

UNIVERSIDADE ESTADUAL DE CAMPINAS. Faculdade de Ciências Médicas. **Treinamento para docentes**: base de dados Dimensions. São Paulo, nov. 2022. Disponível em: <https://www.fcm.unicamp.br/impressa/publicacoes/view/treinamento-para-docentes---base-de-dados-dimensions/14948>. Acesso em: 30 jul. 2025.

8

Rene Faustino Gabriel Junior

ESTUDOS MÉTRICOS DA INFORMAÇÃO NO BRASIL: O USO DA BRAPCI

DOI: 10.31560/pimentacultural/978-85-7221-475-9.8

SUMÁRIO

COMO CITAR:

GABRIEL JUNIOR, Rene Faustino. Estudos métricos da informação no Brasil: o uso da Brapci. In: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.).

Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

As bases de dados bibliográficas vêm passando por um processo de refinamento contínuo que transcende o simples armazenamento referências, títulos, resumos e citações. Atualmente, elas integram funcionalidades analíticas, desde algoritmos de mineração de texto até mecanismos de recomendação baseados em aprendizado de máquina, capazes de mapear, em questão de segundos, os principais autores que estruturam determinado campo e as temáticas emergentes que moldam sua evolução. Essa capacidade sistemática de agrupar e hierarquizar informações não só acelera a descoberta de literatura relevante, mas também permite aos pesquisadores visualizar redes de coautoria, padrões de citação e tendências de publicação ao longo do tempo.

Paralelamente, a incorporação de painéis interativos e dashboards de métricas transformou o modo como se avalia o impacto acadêmico. Ferramentas de análise de citações fornecem indicadores de centralidade e influência de periódicos e pesquisadores, enquanto sistemas de *clustering* identificam automaticamente as lacunas de conhecimento em subáreas específicas. Além disso, a interoperabilidade entre bases de dados, viabilizada por padrões como o OAI-PMH e APIs RESTful, tem permitido o cruzamento de registros e a construção de metarepositórios que ampliam a abrangência da revisão bibliográfica, reduzindo vieses institucionais ou geográficos.

Toda essa oferta de ferramentas de pesquisa eleva o processo investigativo a um novo patamar, permitindo acesso quase instantâneo a resultados quantitativos. Entretanto, a mera exibição

SUMÁRIO

de números descontextualizados, sem uma análise qualitativa aprofundada, sem consideração da conjuntura e do contexto histórico, pouco contribui para os insights e descobertas científicas. É necessário compreender o que esses dados representam e avaliar seu real valor para o avanço do conhecimento. E as metodologias dos Estudos Métricos da Informação (EMI), visam contribuir nessa perspectiva.

OS ESTUDOS MÉTRICOS DA INFORMAÇÃO (EMI)

Os estudos métricos da informação fazem parte de um contexto mais amplo de avaliação e análise da produção científica, inserindo-se na interseção entre a Ciência da Informação e a Sociologia da Ciência (Oliveira, 2018). No cenário internacional, com o crescimento da ciência e da tecnologia, impulsionado por políticas públicas desde a segunda metade do século XX com Bush (1945), exigiu metodologias e abordagens capazes de mensurar, qualificar e interpretar a produção acadêmica e seus impactos (Glanzel; Moed, 2002). Nesse sentido, a aplicação de métodos bibliométricos, cientométricos, informétricos e, mais recentemente, altmétricos tornou-se uma ferramenta para a compreensão do desenvolvimento da ciência (Gouveia, 2013; Araújo, 2015).

Os métodos bibliométricos sofreram influência da matemática e da estatística ao fornecerem os modelos quantitativos e probabilísticos necessários para a análise e interpretação da produção científica (Prichard, 1969). A partir do desenvolvimento de leis como as de Lotka (1926), Bradford (1934) e Zipf (1949), a bibliometria passou a utilizar distribuições estatísticas para modelar a produtividade dos pesquisadores, a dispersão de artigos em periódicos e a frequência de citações (Glanzel, 2003). A Matemática, por meio da teoria

SUMÁRIO

dos grafos e análise de redes, também contribuiu para o estudo de coautorias e citações e cocitações, enquanto a estatística viabilizou a criação de indicadores bibliométricos, como o Fator de Impacto, Índice h e os índices de colaboração científica.

A institucionalização dos EMI no Brasil ocorreu paralelamente à consolidação do ensino e da pesquisa em Ciência da Informação. O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), originalmente IBBD (Instituto Brasileiro de Bibliografia e Documentação), teve um papel incitador ao introduzir estudos bibliométricos com a criação do primeiro programa de Pós-Graduação em Ciência da Informação em 1970.

Os primeiros resultados da aplicação de métodos bibliométricos foram publicados em periódicos brasileiros de Ciência da Informação, destacando-se os trabalhos de Gilda Braga (1973), Laura Maia de Figueiredo (1973) e Elza Lima e Silva Maia (1973) no IBICT, além de Edson Nery da Fonseca (1973) na UnB. Esses estudos pioneiros fundamentavam-se em abordagens quantitativas.

Nas décadas seguintes, e com crescimento do número de pesquisadores, os estudos sobre avaliação da produção científica passaram a se expandir e a integrar diferentes abordagens, tanto quantitativas quanto qualitativas, principalmente associando os dados a elementos cientométricos e a documentos para justificar ou refutar os indicadores (Gabriel Junior, 2014).

Aliadas ao avanço das Tecnologias de Informação e Comunicação (TIC) proporcionou acesso a grandes bases de dados científicas, ampliando as possibilidades de aplicação dos estudos métricos. A partir dos anos 1990, com a expansão da internet e a popularização das bases indexadoras como Web of Science e a criação da Scopus em 2004, os pesquisadores brasileiros passaram a ter mais fontes de dados para analisar a produção qualificada e mapear redes de colaboração, analisar padrões de citação e identificar tendências emergentes em diversas áreas do conhecimento.

SUMÁRIO

Também houve um crescimento do número de artigos produzidos na área de CI (Arboit; Bufrem, 2011) e estudos de compreensão dessa evolução como os estudos de Araujo, (2009); Ferreira (2010); e Oliveira; Alves; Grácio, 2012). Esse movimento foi acompanhado pelo surgimento de eventos científicos e grupos de pesquisa voltados para a temática, como o Grupo de Trabalho 7 da ANCIB (Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação) e os Encontros Brasileiros de Bibliometria e Cientometria (EBBC) (Oliveira, 2018).

Os EMI fornecem um arcabouço teórico e metodológico para a análise da ciência, estruturando um ecossistema de pesquisas baseadas em técnicas bibliométricas, cientométricas, informétricas e, mais recentemente, altmétricas, aplicadas à mensuração da produção científica e da disseminação da informação (Gouveia, 2013).

A Análise de Domínio (AD), proposta por Hjørland e Albrechtsen (1995), influenciou os EMI no Brasil ao proporcionar uma abordagem contextualizada que vai além da contagem de publicações e citações. Essa perspectiva permitiu mapear as estruturas conceituais e epistemológicas da área, identificar tendências emergentes, compreender redes de colaboração, pesquisadores e instituições influentes, além de possibilitar uma compreensão mais profunda da dinâmica da ciência. Diferente de abordagens puramente quantitativas, a AD incorpora dimensões históricas, sociais e cognitivas, permitindo uma análise da conjuntura do desenvolvimento dos EMI.

Compreende-se que, no Brasil, especialmente na área de Ciência da Informação, a primeira vertente ainda predomina, concentrando-se no aprimoramento de modelos estatísticos e ferramentas analíticas; entretanto, observa-se um crescimento significativo das aplicações direcionadas às disciplinas científicas, com estudos que exploram redes temáticas e padrões de colaboração, bem como um interesse emergente pela utilização de indicadores bibliométricos em políticas acadêmicas e na gestão de instituições de pesquisa.

SUMÁRIO

A consolidação dos Estruturas de Métricas e Indicadores (EMI) no Brasil reflete a urgência de adotar metodologias rigorosas capazes de mensurar qualidade, impacto e relevância de forma transparente e confiável. A partir de iniciativas como foi proposto no VIII EBBC–Alagoas (2022) na adoção de métricas responsáveis, com a padronização dos processos de coleta e análise de dados bibliométricos, permitindo comparações mais consistentes entre instituições, áreas de conhecimento e períodos temporais. Além disso, a incorporação de indicadores compostos, que combinam métricas de citações, coautorias e fatores contextuais como colaboração internacional, passou a oferecer uma visão mais holística do desempenho acadêmico, porém mantendo os clássicos como base.

OS CLÁSSICOS DOS ESTUDOS MÉTRICOS DA INFORMAÇÃO (EMI)

Para compreender melhor as aplicações, limitações e o uso dos Estudos Métricos da Informação (EMI), é fundamental conhecer três leis consideradas pilares da bibliometria. Essas leis ajudam a entender os padrões de produção, disseminação e uso da informação científica, sendo amplamente aplicadas em investigações sobre a comunicação científica, avaliação de periódicos, impacto da pesquisa e organização do conhecimento. Elas explicam fenômenos como a produtividade de autores, a frequência de citações e a dispersão da literatura científica, revelando padrões de distribuição, crescimento e uso da informação. As três principais leis da bibliometria são apresentadas a seguir.

Lei de Lotka (Lei da Produtividade dos Autores) que analisa a produtividade dos autores em um determinado campo do conhecimento. A lei estabelece que um pequeno número de autores é responsável por uma grande quantidade de publicações, enquanto a

SUMÁRIO

maioria dos autores publica poucos trabalhos. Ela segue um padrão de distribuição inversamente proporcional, onde a produtividade dos autores diminui exponencialmente à medida que aumenta o número de publicações. Isso significa que, se há 100 autores que publicam um artigo, haverá aproximadamente 25 que publicam dois artigos, 11 que publicam três, e assim por diante. Essa lei é importante para estudos de produtividade científica e para identificar pesquisadores altamente produtivos em uma área.

Lei de Bradford (Lei da Dispersão da Informação) que analisa a dispersão dos artigos científicos entre periódicos. Bradford observou que a literatura científica sobre um tema específico não está distribuída uniformemente entre os periódicos. Em qualquer área, poucos periódicos concentram a maior parte dos artigos relevantes, enquanto um grande número de periódicos publica poucos artigos sobre o tema. Utilizada para otimizar a seleção de periódicos em bibliotecas e bases de dados, além de ser aplicada na análise de fontes para revisões bibliográficas. Essa distribuição é organizada em três zonas: Zona 1 - periódicos altamente relevantes concentra um grande volume de artigos; Zona 2 - um número intermediário de periódicos publica menos artigos sobre o tema; e Zona 3 - muitos periódicos contêm apenas alguns artigos sobre o assunto.

E a Lei de Zipf (Lei da Frequência das Palavras), que estabelece a relação entre a frequência e a classificação das palavras em um texto ou conjunto de textos científicos. Zipf demonstrou que, em um grande conjunto de documentos, a frequência de uma palavra é inversamente proporcional à sua posição no ranking de uso. Ou seja, a segunda palavra mais usada aparece aproximadamente metade das vezes da mais frequente, a terceira palavra um terço das vezes, e assim por diante. Usada para análise de linguagem natural, indexação de documentos, recuperação da informação e estudos de terminologia.

SUMÁRIO

Os indicadores bibliométricos podem ser classificados em três grandes grupos (Spinack, 1998; Moed, 2005): (i) produtividade, que quantifica o volume de publicações de um autor, instituição ou país; (ii) impacto, que avalia a influência dessas publicações por meio de citações (como o fator de impacto de periódicos e o índice h de autores); e (iii) colaboração, que analisa as redes de coautoria e parcerias institucionais ou internacionais, evidenciando a estrutura e a densidade das redes científicas.

Ainda na avaliação, um outro indicador importante nas pesquisas de produção é a Lei de Price, proposta por Derek J. de Solla Price em 1963, onde afirma que, em qualquer campo científico ou coletivo de pesquisadores, metade da produção total é gerada por um número de autores que corresponde à raiz quadrada do número total de autores envolvidos. Em outras palavras, se existem N autores atuando numa área, apenas $\sqrt{N}$ deles serão responsáveis por 50 % dos artigos publicados. Essa “regra da meia-produção” evidencia a forte concentração de produtividade em um pequeno grupo de pesquisadores.

Em 1969, no artigo “Statistical Bibliography or Bibliometrics?”, Alan Pritchard introduziu o termo bibliometrics, bibliometria em português, propondo a substituição da expressão então vigente “estatística bibliográfica” por “bibliometria”, entendida como a aplicação de métodos matemáticos e estatísticos a livros e outros meios de comunicação.

De acordo com Glänzel (2003), os estudos bibliométricos, ou EMI, apresentam-se em três grandes vertentes de aplicação: (i) Bibliometria para bibliometristas, destinada ao desenvolvimento metodológico e à investigação fundamental das técnicas e modelos estatísticos empregados na análise de produção e citação científica; (ii) Bibliometria para disciplinas científicas, cujo foco está na extensão dos métodos bibliométricos às diversas áreas do conhecimento, permitindo aos pesquisadores mapear perfis de publicação, redes de

SUMÁRIO

coautoria e dinâmicas temáticas específicas de seu campo de atuação; e (iii) Bibliometria para políticas e gestão da ciência, voltada à avaliação institucional e regional, onde indicadores agregados subsidiavam decisões de fomento, comparações entre países ou instituições e monitoramento de programas de pesquisa.

No contexto dos indicadores de avaliação, sobressaem o Fator de Impacto (FI) e o índice h . Contudo, é importante assinalar diversas críticas ao seu emprego, sobretudo no que concerne à dependência das bases de dados subjacentes, que apresentam vieses linguísticos e geográficos, como ocorre em plataformas dominantes, tais como Web of Science e Scopus, limitando a visibilidade de produções originárias de regiões periféricas ou publicadas em periódicos não indexados nesses grandes agregadores.

O Fator de Impacto (FI), conforme proposto por Garfield (1965), corresponde à média aritmética de citações recebidas em um determinado ano pelos artigos publicados em um periódico nos dois anos imediatamente anteriores. Em sua concepção, o FI oferece um indicador quantitativo da influência de uma revista científica, permitindo comparar periódicos segundo o impacto médio de seus artigos e subsidiando decisões editoriais e de avaliação da produção acadêmica – embora Garfield também ressalte limitações inerentes, como a variação entre disciplinas e o efeito de autocitações (GARFIELD, 1965).

O índice h , proposto por Hirsch (2005), é definido como o maior valor h tal que o pesquisador tenha publicado h artigos que receberam, cada um, pelo menos h citações. Esse indicador combina produtividade e impacto acadêmico de forma integrada, proporcionando uma medida mais robusta do desempenho científico individual em comparação a métricas que consideram isoladamente número de publicações ou total de citações.

SUMÁRIO

Essa dependência de indicadores quantitativos consolidados em bases dominantes não só reforça a concentração de visibilidade nas mesmas publicações e pesquisadores, mas também alimenta o Efeito Mateus descrito por Merton (1968), pelo qual periódicos e autores já prestigiados atraem ainda mais citações e atenção. Assim, ao combinar o Fator de Impacto, o índice h e o fenômeno do Efeito Mateus, percebe-se um ciclo de retroalimentação em que métricas originalmente concebidas para mensurar influência acabam por legitimar e perpetuar a desigualdade na distribuição de reconhecimento acadêmico, enfatizando a necessidade de abordagens complementares que considerem tanto aspectos qualitativos quanto contextuais na avaliação da produção científica.

Ao modelar citações como redes dirigidas, Price (1965) estabeleceu as bases para os estudos de acoplamento bibliográfico, proposto por Kessler (1963) para quantificar a similaridade entre documentos que compartilham referências, e de co-citação, formulado por Small (1973) para identificar relações temáticas entre artigos citados em conjunto, consolidando o arcabouço metodológico necessária ao mapeamento da estrutura e evolução do conhecimento científico.

O Índice de Price corresponde à porcentagem de referências com até cinco anos de publicação em relação ao total de referências de um documento ou de um conjunto documental, sendo um indicador da atualidade e relevância das fontes consultadas. Por exemplo, imagine um artigo de revisão em Ciência da Informação que reúna 50 referências, das quais 30 foram publicadas nos últimos cinco anos. Nesse caso, o Índice de Price será de 60 %, indicando que a maioria das fontes utilizadas é recente e confere maior atualidade ao levantamento bibliográfico.

No âmbito da Análise de Redes Sociais (ARS), Newman (2001) foi pioneiro ao empregar métricas da teoria dos grafos, como grau, caminhos mínimos e coeficiente de agrupamento, para mapear redes de coautoria em periódicos científicos, estabelecendo bases

SUMÁRIO

para a caracterização quantitativa de conectividade e centralidade em comunidades acadêmicas, conforme proposto por Freeman (1979) na definição e comparação formal de medidas de centralidade (grau, proximidade e intermediação) para a identificação de posições estruturais nodais.

Os estudos de indicadores no contexto brasileiro e latino-americano beneficiam-se da Brapci como ambiente de aplicação e validação, pois ela reúne metadados e textos completos de periódicos e eventos regionais. Esse corpus diversificado minimiza vieses geográficos e linguísticos das bases comerciais e retrata com maior fidelidade as dinâmicas científicas periféricas, promovendo uma avaliação mais inclusiva e representativa da produção acadêmica na região.

A BASE DE DADOS EM CIÊNCIA DA INFORMAÇÃO (BRAPCI)

Criada em 2006 como Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação, a Brapci reuniu, organizou e disponibilizou a produção científica brasileira na área. Inicialmente focada no mapeamento de artigos publicados, ela logo se consolidou como ferramenta de apoio à pesquisa e à disseminação do conhecimento (Bufrem *et al.*, 2010).

Ao longo dos anos, a plataforma passou a oferecer não apenas referências, mas também, sempre que possível, textos completos, ampliando a visibilidade da literatura nacional e preservando sua memória, passando a se chamar de Base de Dados em Ciência da Informação. Sua arquitetura evoluiu com tecnologias de organização da informação, classificação temática e indexação, permitindo análises bibliométricas detalhadas, mapeamentos temáticos e estudos de redes de coautoria e cocitação.

SUMÁRIO

Desenvolvida e mantida por pesquisadores brasileiros, a Brapci destaca-se pela ampla cobertura de periódicos latino-americanos em português e espanhol. Além da recuperação de metadados, ela disponibiliza ferramentas para estudos métricos — como análises de produção científica, redes colaborativas e padrões de citação — auxiliando avaliações acadêmicas e institucionais e identificando tendências de pesquisa.

Com buscas avançadas por autor, título, periódico, palavras-chave e ano, e integração a sistemas de referência e de análise de dados, a plataforma oferece uma interface intuitiva e atualizações constantes. Esses recursos permitem aos usuários conduzir estudos aprofundados sobre a evolução do conhecimento na Ciência da Informação.

No contexto da ciência aberta, a Brapci reforça seu compromisso com a democratização do saber, reunindo e organizando dados de forma acessível. Entretanto, desafios persistem, como ampliar a integração com outras áreas do conhecimento, fortalecer bases nacionais e aprimorar metodologias capazes de refletir as novas dinâmicas da comunicação científica.

AS APLICAÇÕES DOS ESTUDOS MÉTRICOS DA INFORMAÇÃO NA BRAPCI

A qualidade do resultado de um EMI depende muito da estratégia de busca empregada na recuperação. Para se ter uma recuperação eficiente é necessário estabelecer uma boa estratégia de busca. Para isso é importante entender a lógica de funcionamento dos motores de busca.

SUMÁRIO

Uma estratégia de busca bem estruturada precisa assegurar a recuperação eficiente de informação. Isso exige a definição clara do tema central e a seleção dos termos mais representativos, incluindo variantes e sinônimos. Recomenda-se, portanto, o uso de vocabulários controlados ou tesouros, capazes de mapear especificidades conceituais. Por exemplo, ao investigar “Inteligência Artificial”, além do termo genérico, deve-se contemplar conceitos correlatos, deep learning, machine learning, visão computacional, ChatGPT, entre outros, para ampliar a abrangência e a precisão dos resultados.

Também é importante destacar que, em muitas pesquisas, os motores de busca automaticamente separam as palavras inseridas, utilizando o operador booleano AND entre elas. Essa estratégia visa recuperar o máximo possível de resultados, mas pode retornar documentos que não tratam especificamente do assunto desejado. Por isso, recomenda-se o uso de aspas para delimitar palavras compostas, garantindo maior precisão na recuperação, por exemplo, ao buscar por “Inteligência Artificial”, a base considerará a expressão exata, e não os termos de forma isolada.

Em algumas bases de dados, como a Brapci, é possível realizar buscas booleanas ou avançadas, utilizando operadores como OR (para ampliar a busca) ou NOT (para excluir termos indesejados). Como ilustrado na Figura, uma consulta pode ser construída para recuperar documentos que abordem Inteligência Artificial e ChatGPT, mas que excluam aqueles relacionados a Desinformação ou Revisão Sistemática.

Quadro 1 - Estratégia de busca booleana

“Inteligência artificial” AND chatgpt NOT
(desinformação OR “revisão sistemática”)

Fonte: elaborado pelo autor, 2025.

SUMÁRIO

O exemplo apresentado na Figura 1 ilustra uma estratégia de busca elaborada para recuperar trabalhos que contenham simultaneamente as expressões "Inteligência Artificial" e "ChatGPT", excluindo, no entanto, aqueles que abordam os temas *desinformação* e *revisão sistemática*. Essa formulação de busca, baseada em operadores booleanos, é um componente fundamental da metodologia da pesquisa e deve ser devidamente registrada e descrita. Sua explicitação permite a reprodutibilidade do estudo, seja com o objetivo de ampliá-lo, atualizá-lo em novos contextos ou realizar monitoramentos contínuos sobre o tema.

Outro fator importante a ser considerado nas pesquisas é o recorte temporal, que pode ser sincrônico ou diacrônico. A abordagem diacrônica permite uma visão global e evolutiva ao longo de todo o período analisado. Por exemplo, ao se estudar a produção científica da Ciência da Informação no Brasil entre 1972 e 2024, contempla-se desde as primeiras publicações registradas em periódicos da área (1972) até o ano mais recente com dados completos (2024).

Já a análise sincrônica, voltada para comparações entre períodos distintos, exige maior rigor metodológico. De acordo com os princípios estabelecidos na clássica obra de Solla Price (1987), "*Little Science, Bi Science*", esse tipo de análise deve abranger, no mínimo, cinco anos por intervalo, de modo a garantir robustez estatística. Quando há mais de um período de comparação, os recortes devem ter a mesma duração temporal. Por exemplo: ao comparar a produção científica das décadas de 1990 e 2010, é recomendável estabelecer os recortes de 1990 a 1999 e de 2010 a 2019. Não se deve incluir algum período não completo, como por exemplo a década de 2020 (pois estamos em 2025), ou quando o ano corrente quando analisado por ano.

SUMÁRIO

ANÁLISE DE PRODUÇÃO

Um aspecto importante a ser considerado na análise de publicações recentes, especialmente aquelas dos últimos três anos, é que alguns periódicos enfrentam dificuldades em manter a regularidade de suas edições, o que pode resultar em atrasos na publicação de fascículos. Com isso, é comum que edições retroativas sejam publicadas apenas no ano corrente, o que pode gerar distorções nas análises de produção anual. Esse descompasso pode fazer com que a produção de anos anteriores continue a crescer artificialmente ao longo do tempo, comprometendo a precisão temporal dos dados. Além disso, tal inconsistência pode ocasionar problemas em análises de citações, como a ocorrência de referências a artigos cuja data de publicação formal é posterior à data de publicação do artigo.

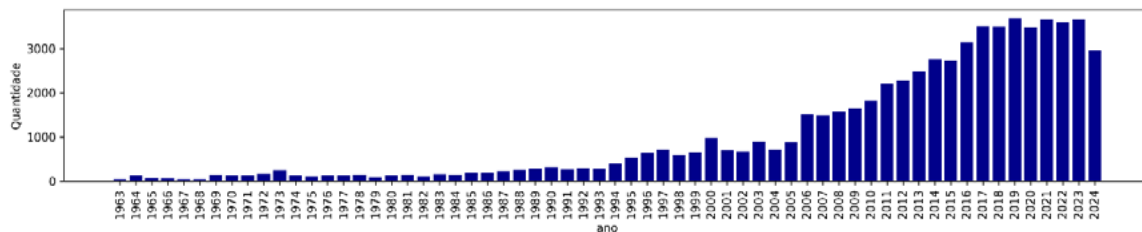
Algumas revistas recorrem à prática de atribuir datas de publicação futuras a artigos já disponíveis online, estendendo artificialmente sua janela de citação. Embora isso possa aumentar o número de citações dentro do período formal de vigência, tal estratégia distorce a análise cronológica da produção científica e compromete as estatísticas anuais de publicação, exigindo cautela redobrada em estudos bibliométricos e avaliações acadêmicas.

É frequente, em análises temáticas de bases de dados, afirmar que a produção de um determinado tema vem crescendo com base em gráficos de séries temporais. No entanto, como a produção anual total em todas as áreas do conhecimento também está aumentando, torna-se necessário comparar o crescimento específico de um tema com o crescimento global da publicação científica. Sem essa referência, por exemplo, calculando a participação percentual de um tema em relação ao total anual de publicações, não é possível validar se o avanço observado decorre de um interesse temático real ou apenas reflete a expansão geral do corpus acadêmico.

SUMÁRIO

Para ilustrar, foram elaborados os Gráficos 1, 2 e 3. O Gráfico 1 apresenta o total de trabalhos indexados na Brapci entre 1963 e 2024 (diacrônico) (Gabriel Junior, 2025), evidenciando o crescimento da produção ao longo do tempo e sua estabilização a partir de 2016. Já o Gráfico 2 mostra o número absoluto de publicações sobre um tema específico, dando a impressão de expansão; porém, quando comparado ao Gráfico 1, fica claro que essa temática tem, na verdade, perdido participação relativa, uma vez que seu ritmo de crescimento não acompanha o aumento da produção total. Por fim, o Gráfico 3 corrige essa percepção ao exibir a evolução percentual do tema em relação ao total de artigos, confirmando seu real avanço. Dessa forma, torna-se evidente a necessidade de empregar indicadores comparativos para evitar conclusões precipitadas. Para facilitar essa comparação pode-se utilizar o coeficiente de Pearson (1896), para medir essas relações. Ressalta-se também a importância da escala utilizada em cada gráfico: enquanto o Gráfico 1 abrange até 3 000 trabalhos indexados, o Gráfico 2 concentra-se em um máximo de 60 publicações temáticas e o Gráfico 3 em torno de 250 registros proporcionais ao total, garantindo a adequada visualização dos diferentes níveis de agregação de dados.

Gráfico 1 - Produção total da Brapci de artigos, trabalhos de eventos, livros e capítulos



SUMÁRIO

Gráfico 2 - Produção TEMA1 da Brapci

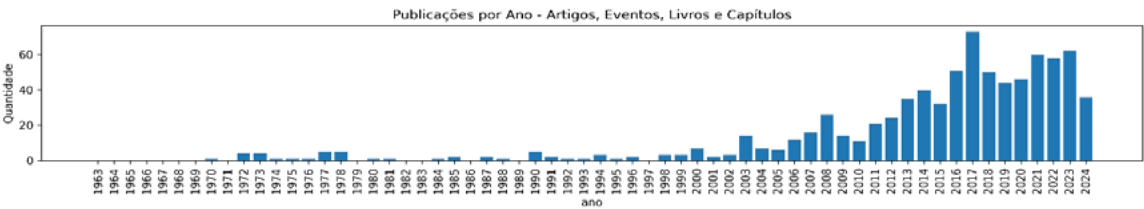
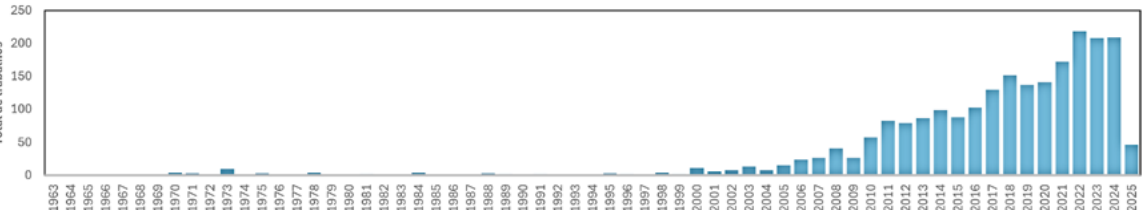


Gráfico 3 - Produção TEMA2 da Brapci



Fonte: Dados da pesquisa (2025).

Outra forma de aprimorar a qualidade das análises nos EMI é por meio da aplicação de técnicas de análise multivariada. Esse conjunto de métodos estatísticos permite examinar simultaneamente múltiplas variáveis, identificando padrões ocultos, relações complexas e estruturas subjacentes nos dados. Com a análise multivariada, torna-se possível, por exemplo, agrupar documentos ou autores com características semelhantes (análise de cluster), reduzir a dimensionalidade dos dados para facilitar a visualização e interpretação (análise fatorial ou de componentes principais), além de explorar correlações entre variáveis que influenciam o comportamento da produção científica.

Nessa mesma interface, a Brapci oferece recursos de exportação em múltiplos formatos, como CSV, Excel, RIS, BibTeX, entre outros, facilitando a análise posterior em softwares bibliométricos ou gerenciadores de referência. Além disso, é possível acessar um painel com indicadores bibliométricos, que apresenta informações

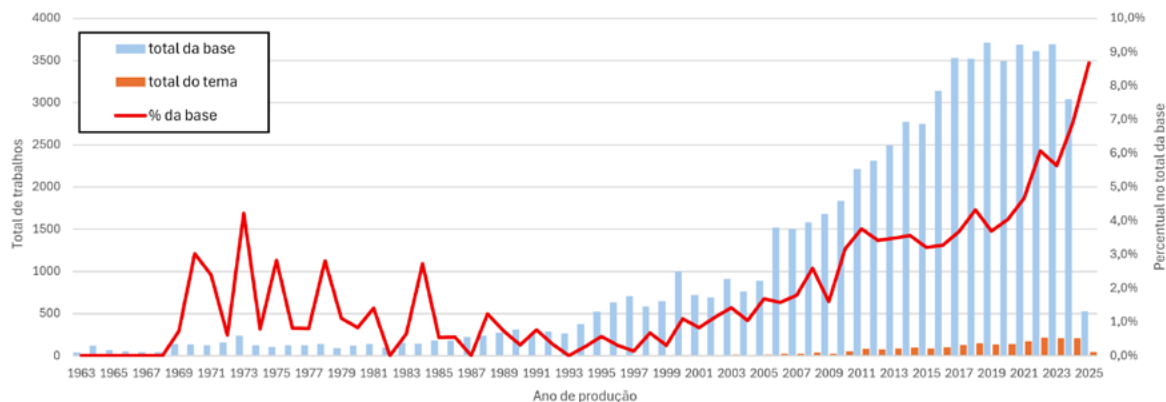
SUMÁRIO

como número de documentos por ano, autores mais produtivos, palavras-chaves mais representadas e outras métricas relevantes para análises comparativas.

Como resultado, foram recuperados 2.253 registros, dos quais 1.795 são artigos de periódicos, 1.142 (63,6 %) nacionais e 653 (36,4 %) internacionais, além de 450 trabalhos em eventos, 2 livros e 6 capítulos de livros. Para facilitar a análise preliminar desses dados, utiliza-se o recurso “Painel”, que apresenta uma visualização dos principais indicadores de produção, apresentado no Gráfico 4.

Observa-se, em análise diacrônica, que a temática emergiu no final da década de 1960 com participação ainda reduzida, alcançando mais de 4 % do total de trabalhos na década de 1970. Nas décadas de 1980 e 1990, sua representatividade caiu para abaixo de 1 %, indicando diminuição de interesse. A partir dos anos 2000, porém, o tema retoma relevância contínua, projetando-se em torno de 9 % das publicações até 2025.

Gráfico 4 – Percentual da produção sobre Bibliometria na produção total (n=2.253)



Fonte: Dados da pesquisa (2025).

SUMÁRIO

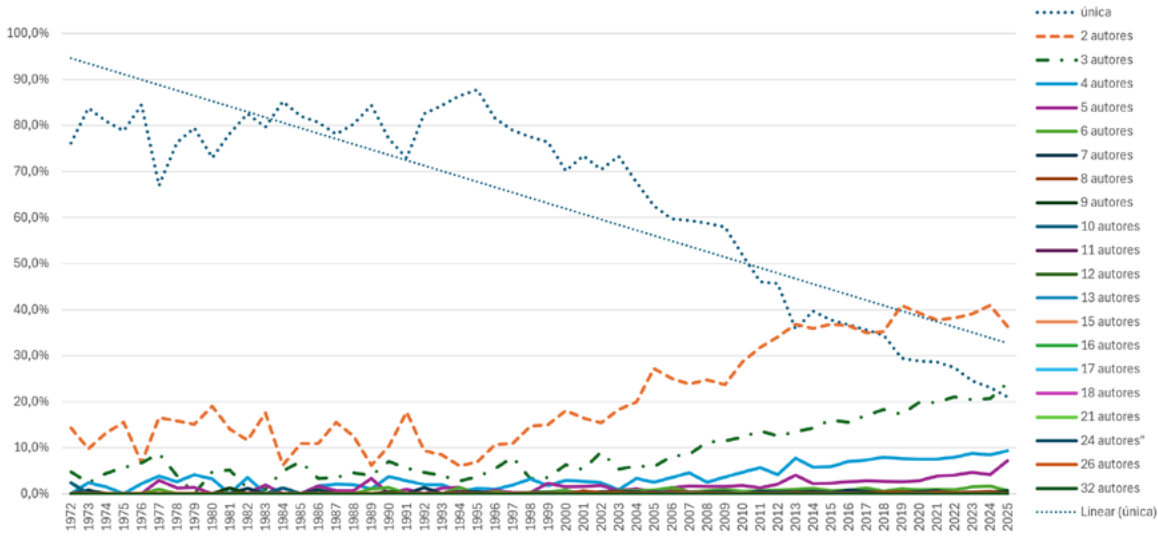
A análise foi realizada com base nos dados extraídos da plataforma, o que minimiza possíveis erros de interpretação. Para fins comparativos, a evolução temática é apresentada quantitativamente no Gráfico 3. Vale ressaltar que, embora 2025 ainda esteja em curso, seu valor não distorce o indicador temático, pois este é calculado de forma proporcional. Essa análise diacrônica só se torna viável graças ao acesso detalhado aos dados fornecidos pela plataforma. Contudo, quando o volume de dados é reduzido, os resultados podem ficar distorcidos, como evidencia o Gráfico 4: a linha de representatividade somente se estabiliza a partir de 1994, quando a quantidade de informações se torna suficientemente expressiva.

ANÁLISE DE AUTORIA E COAUTORIA

As análises de autoria única e de coautoria têm sido amplamente investigadas na literatura da Ciência da Informação. Por exemplo, Queiroz e Vilan Filho (2019) dedicaram-se a examinar em detalhe as práticas de autoria única, identificando fatores que levam o pesquisador a assinar trabalhos de forma isolada. Em um panorama mais amplo, Bufrem, Gabriel Júnior e Gonçalves (2010) empregaram métodos bibliométricos para mapear padrões de colaboração e redes de coautoria em periódicos, livros e eventos brasileiros, revelando tendências históricas e disciplinares. Além disso, estudos focalizados em contextos institucionais, como o de Cajazeira e Silva (2021), que analisou a produção científica da Universidade Federal do Piauí (UFPI), ofereceram subsídios qualitativos sobre as dinâmicas locais de cooperação entre professores e alunos.

SUMÁRIO

Gráfico 5 – Percentual da autoria única e coautorias em Revistas Brasileiras (1972-2025)



Fonte: Dados da pesquisa (2025).

O Gráfico 5 ilustra a evolução da composição autoral dos artigos em Ciência da Informação no Brasil a partir de 1972, apresentando, em termos percentuais, a distribuição entre autoria única, dupla e tripla. Observa-se uma queda contínua na proporção de trabalhos de autoria individual, enquanto a coautoria com dois autores estabilizou-se em torno de 40 % do total. Paralelamente, a participação de três autores vem crescendo de modo consistente, alcançando atualmente cerca de 20 %. Esses padrões sugerem uma transição gradual para modos de pesquisa mais colaborativos, possivelmente impulsionados pela complexidade crescente dos temas investigados e pelo incentivo institucional a parcerias interdisciplinares.

Outro tipo de análise que pode ser realizada com os autores refere-se , e referente a ARS, que será abordada um pouco mais adiante.

SUMÁRIO

ANÁLISE DE PERIÓDICOS

No Brasil, é difícil diferenciar um periódico científico (acadêmico) de um profissional, pois muitas revistas profissionais também publicam reflexões aprofundadas, além dos tradicionais relatos de experiência. A Brapci indexa revistas brasileiras e de países cujo idioma oficial seja o português ou o espanhol. Ao processar as informações, a Brapci enriquece seus metadados, padronizando a entrada de autores, coletando palavras-chave do texto completo e realizando uma série de verificações e validações dos documentos indexados. Esse trabalho é executado continuamente por robôs treinados com regras de validação; outros robôs utilizam inteligência artificial, por exemplo, na tradução de títulos, resumos e palavras-chave para três idiomas: português, inglês e espanhol.

Ao analisar periódicos latino-americanos, é preciso considerar seus objetivos, que se dividem em três categorias: os acadêmicos (ou de pesquisa), que veiculam estudos de pós-graduação com resultados experimentais ou teóricos; os profissionais, voltados à resolução prática de problemas da área; e os de divulgação científica, que traduzem a linguagem acadêmica em termos acessíveis a um público leigo.

Em 2025, a Brapci reúne 79 periódicos brasileiros, 25 periódicos estrangeiros e 8 anais de eventos. O Gráfico 6 traça a trajetória de títulos que foram incorporados ao acervo ao longo das décadas. Logo nos primeiros registros (1963 e 1969) aparecem títulos estrangeiros isolados, mas é só a partir de 1972 que surge o primeiro periódico brasileiro vinculado ao Ibict/UFRJ, marcando o início de uma lenta expansão nacional (3 revistas em 1972 e 3 em 1973). Nas décadas de 1980 e início de 1990, a média fica em torno de 1–2 entradas por ano, refletindo uma fase de consolidação editorial.

Já a virada dos anos 2000 coincide com o fortalecimento dos programas de pós-graduação e o início dos periódicos digitais: em 1999, por exemplo, entram 4 títulos brasileiros e 2 estrangeiros (total de 6), e em

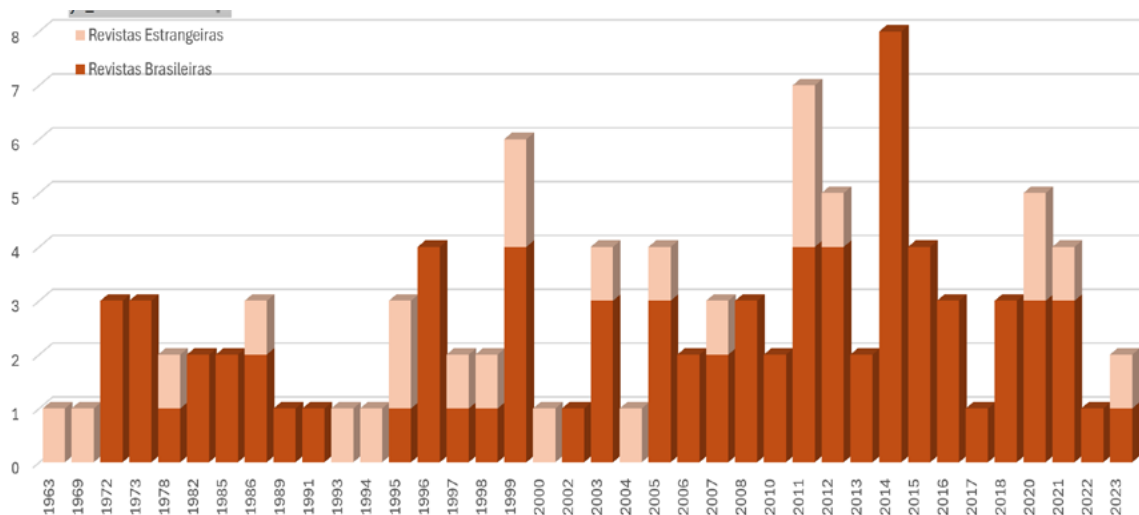
SUMÁRIO

2003 e 2005 surgem 3 periódicos nacionais a cada ano. Destacam-se os picos de 2011 (4 revistas brasileiras e 3 estrangeiras, totalizando 7) e 2012 (4 BR + 1 EXT = 5), período em que a CAPES incrementou a avaliação de periódicos e estimulou novas publicações.

O ápice, contudo, ocorre em 2014, com a incorporação de 8 títulos brasileiros, a maior entrada anual registrada, sinalizando o amadurecimento do campo e o alinhamento com políticas de fomento à pesquisa. Após este pico, observa-se um leve decréscimo, mas ainda significativo, com médias entre 2 e 4 novos periódicos nacionais por ano até 2023.

Neste contexto, pode-se inferir que os periódicos brasileiros passaram por três fases: Iniciação (1963–1990): poucos títulos, observando-se maior densidade nos anos de criação dos primeiros programas de pós-graduação. Expansão moderada (1991–2009): entradas regulares de 1 a 4 periódicos nacionais, com o complemento de alguns títulos estrangeiros. Consolidação (2010–2025): reflexo direto das políticas de avaliação da CAPES e do aumento dos investimentos acadêmicos, culminando no pico de 2014.

Gráfico 6 – Criação de Revistas Indexadas na Brapci (1963-2023)



Fonte: Dados da pesquisa (2025).

SUMÁRIO

INDICADORES DE PRODUÇÃO
NO PAINEL DA BRAPCI

Na Brapci, ao executar uma consulta, o usuário pode clicar em “Selecionar” para agrupar os registros de seu interesse. Imediatamente, o botão passa a exibir “Ver xxx itens selecionados”, permitindo que o pesquisador visualize e analise rapidamente esse subconjunto. Um novo clique nesse botão leva diretamente à tela de exportação dos dados selecionados, como mostrado na Figura 1. Como já mencionado, é possível exportar os dados para vários formatos. Além do sistema disponibilizar um resumo da tipologia das fontes.

Para usuários cadastrados e logados na Brapci, não há restrição quanto ao volume de registros que podem ser exportados em uma única operação: é possível gerar arquivos completos nos formatos CSV, RIS, BibTeX, EndNote ou XML conforme a necessidade da pesquisa. Já os usuários anônimos (que não realizaram login) têm um limite máximo de 2.000 registros por exportação, garantindo desempenho adequado do sistema e evitando sobrecarga no servidor. Caso seja necessário obter um conjunto maior de resultados, o usuário não autenticado deve refinar a busca em rodadas sucessivas ou optar por criar uma conta na plataforma, liberando o acesso ilimitado às ferramentas de exportação.

Figura 1 – Resumo da Seleção de Trabalhos



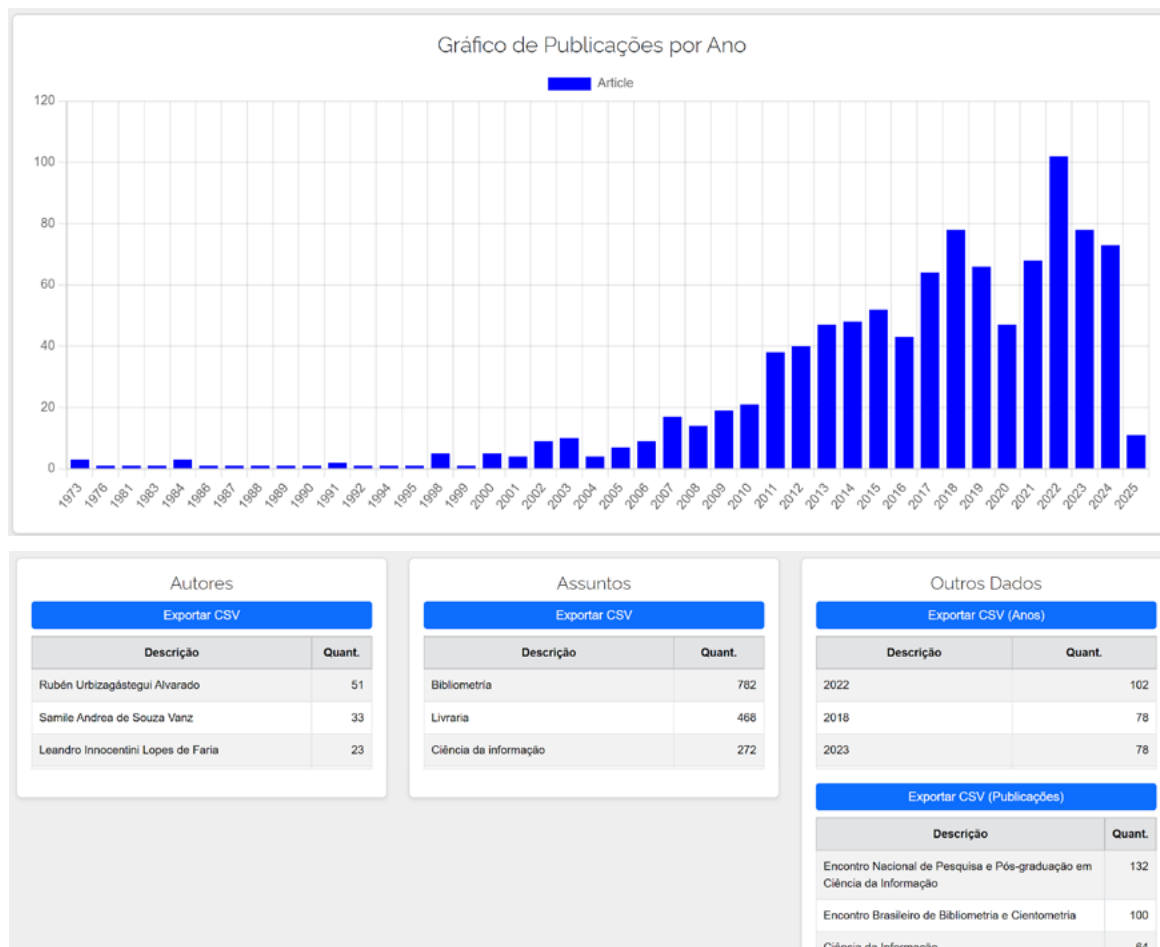
Fonte: Autor (2025).

No canto direito da página, há um botão identificado como “Painel”, que abre uma área dedicada a indicadores de produção já calculados. Ao acionar esse botão, o pesquisador tem acesso

SUMÁRIO

imediatamente a relatórios pré-processados – como distribuição temporal de publicações, redes de coautoria, métricas de impacto e frequência de palavras-chave – organizados em arquivos prontos para visualização ou exportação. Esse recurso centraliza em um único espaço os principais indicadores bibliométricos, agilizando a análise e permitindo uma interpretação rápida dos resultados sem a necessidade de processamento manual adicional.

Gráfico 7 – Painel de exportação de dados na Brapci



Fonte: Dados da pesquisa (2025).

SUMÁRIO

Na avaliação de indicadores de produção, é necessário ir além dos números e incorporar, por exemplo, a Análise de Domínio (Hjørland; Albrechtsen, 1995) para contextualizar e interpretar os dados. Em vez de apenas quantificar a frequência de publicações ou citações, essa abordagem permite, por exemplo, identificar os autores mais prolíficos em determinado tema e, simultaneamente, mapear a instituição a que estão vinculados, sua formação acadêmica e seu protagonismo no campo. Ao cruzar esses elementos qualitativos com os indicadores bibliométricos, o pesquisador pode conjecturar sobre as redes de influência, entender como práticas institucionais e trajetórias pessoais moldam o desenvolvimento do conhecimento e apontar potenciais lacunas ou vieses no domínio estudado. Dessa forma, a combinação de métricas e Análise de Domínio oferece uma visão mais rica e crítica da dinâmica científica.

ANÁLISE DE REDES SOCIAIS

A Análise de Redes Sociais (ARS) aplica conceitos e métricas da teoria dos grafos para mapear e quantificar as relações de colaboração e citação entre autores, instituições ou documentos. Por meio de medidas como grau, proximidade, intermediação e coeficiente de agrupamento, é possível identificar os nós mais centrais, as pontes estruturais que conectam diferentes comunidades e os agrupamentos temáticos emergentes (Freeman, 1979; Newman, 2001). Além disso, algoritmos de detecção de comunidades, como o método de Louvain ou o de Girvan–Newman, permitem revelar clusters de cooperação, apontando redes de pesquisa consolidadas e potenciais lacunas de integração entre grupos.

No ambiente da Brapci, a ARS é operacionalizada por meio da exportação de grafos no formato .net, compatível com ferramentas como Gephi, Pajek, VozViewVOSviewer ou UCInet. Após a

SUMÁRIO

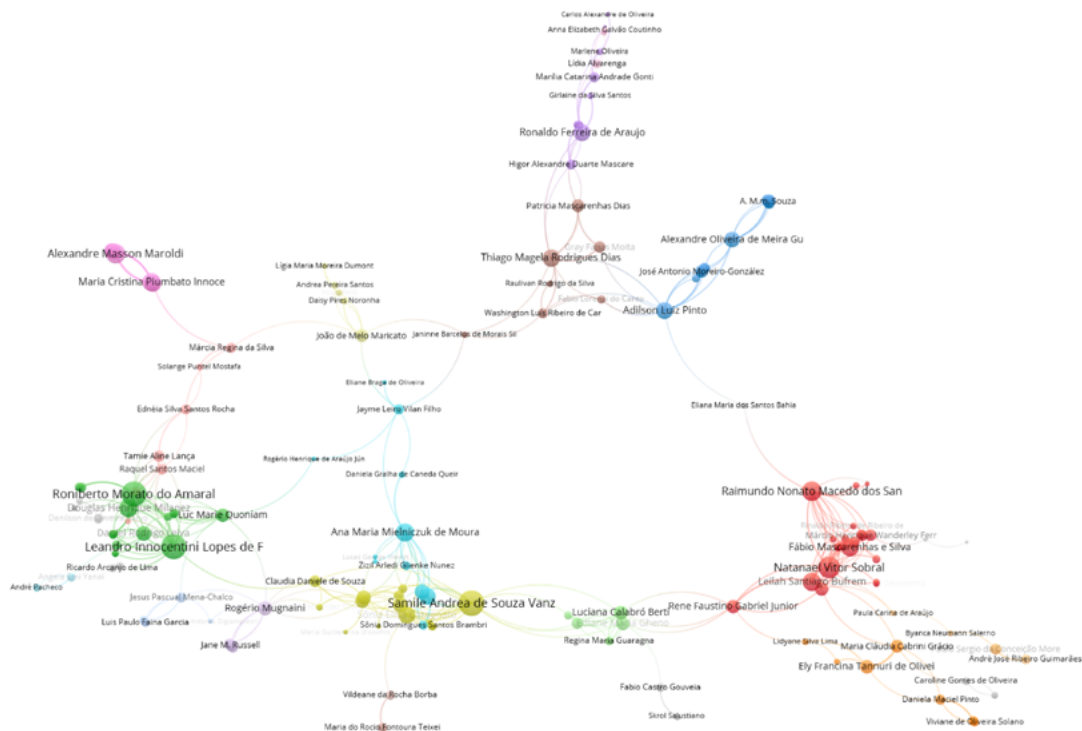
conversão, o pesquisador pode aplicar filtros de nós (por instituição, área temática, período) e calcular, de forma automatizada, indicadores de centralidade e modularidade. Essa funcionalidade não apenas agiliza a construção de mapas de coautoria e de cocitação, mas também possibilita análises comparativas entre subdomínios ou programas de pós-graduação, evidenciando, por exemplo, quais grupos institucionais atuam como hubs de difusão do conhecimento.

Da mesma forma na análise de produção, para enriquecer os resultados quantitativos da ARS, recomenda-se integrá-los à Análise de Domínio (Hjørland & Albrechtsen, 1995), cruzando dados topológicos com informações qualitativas sobre formação, filiação institucional e trajetória acadêmica dos atores. Dessa forma, além de mensurar quem ocupa posições centrais na rede, é possível conjecturar sobre seu papel na consolidação de temas específicos, os fluxos de inovação entre subáreas e a influência de políticas institucionais na dinâmica colaborativa. Essa abordagem híbrida oferece uma visão mais crítica e contextualizada das redes científicas, fundamental para subsidiar decisões de fomento e gestão da pesquisa.

Por exemplo, foi realizada uma busca por “bibliometria” na Brapci, resultando em 1 005 trabalhos, que foram exportados em formato XLS. Em seguida, a coluna AUTHORS foi extraída e colada em um arquivo de texto (.TXT), do qual se removeu a primeira linha contendo o cabeçalho. Esse arquivo foi então submetido à ferramenta de conversão bibliográfica da Brapci (acessível a usuários logados), que transforma o TXT em um grafo no formato .NET. Ao final do processo, a plataforma gera o arquivo .NET pronto para análise em softwares de Análise de Redes Sociais. O resultado é apresentado na Figura 7.

SUMÁRIO

Gráfico 8 – Análise de Redes Sociais sobre Bibliometria



Fonte: Dados da pesquisa (2025).

O Gráfico 8, gerado no VozView com parâmetros de atração = 5 e repulsão = 1, evidenciou 26 clusters distintos, revelando as principais coalisões colaborativas na área sobre bibliometria, Evidenciando os principais grupos que trabalham com o tema. Núcleo UFRGS: centrado em Samile Vanz, esse cluster reúne autores ligados à Universidade Federal do Rio Grande do Sul e suas ramificações, indicando a força das linhas de pesquisa em organização da informação e bibliometria naquele programa. Agrupamento UFSCar: liderado por Roniberto Morato do Amaral e Leandro Innocentini

SUMÁRIO

Lopes de Faria, destaca-se pela cooperação interna na Universidade Federal de São Carlos, com ênfase em estudos de redes e classificação. Cluster UFPE: composto por Raimundo Nonato, Fábio Mascarenhas, Natanael Sobral e Leilah Santiago Bufrem, indica uma coesão forte em temas de avaliação de periódicos e métricas de produção, reforçando o protagonismo da Universidade Federal de Pernambuco. Conexão UFSC: onde Adilson Pinto assume papel central, mapeando produção e colaborações na Universidade Federal de Santa Catarina. A parceria UFAL–CEFET-MG: com Ronaldo Araújo e Thiago Magela Dias atuando como pontes, esse cluster demonstra um fluxo de conhecimento entre a Universidade Federal de Alagoas e o CEFET-MG, sinalizando intercâmbio metodológico em ciência da informação aplicada.

Além desses, há diversos clusters menores e pares isolados que refletem grupos emergentes ou pesquisa individual. A existência de 26 agrupamentos mostra a fragmentação e, ao mesmo tempo, a consolidação de redes específicas: grandes hubs institucionais convivem com pequenas equipes temáticas, o que sugere tanto a força de programas bem estabelecidos quanto o surgimento de novos focos de investigação. Essa configuração reforça a importância das políticas de cooperação interinstitucional e aponta direções para futuros estímulos ao intercâmbio acadêmico.

ANÁLISE DOS INDICADORES DE CITAÇÃO

Os estudos de citação esbarra frequentemente na ausência de uma base de dados unificada que reúna e organize as referências de todos os trabalhos científicos. A Brapci tem se empenhado nessa tarefa de coleta e normalização, mas enfrenta o desafio da heterogeneidade de estilos — enquanto algumas revistas adotam a ABNT, outras usam APA, Vancouver ou até padrões proprietários —, o que compromete

SUMÁRIO

a extração automática e o agrupamento consistente das referências. Além disso, a falta de revisão sistemática dos registros de citação aumenta a taxa de erros e omissões, exigindo intervenções manuais pontuais para garantir a qualidade do conjunto de dados.

Em consequência, a análise de indicadores de produção em Brapci costuma ocorrer de forma fragmentada, dependendo muitas vezes da coleta isolada de referências para cada estudo. Convém ressaltar que o que se denomina “análise de citação” consiste, na prática, em examinar as referências efetivamente citadas pelos autores, sem necessariamente quantificar a frequência de menções em toda a literatura. Um exemplo notável é o trabalho de Silveira, Caregnato e Bufrem (2014), que propôs uma classificação das motivações das citações em Ciência da Informação, contribuindo para compreender não apenas o objeto citado, mas também o papel epistemológico atribuído a ele.

Para mitigar essas limitações, as ferramentas da Brapci passaram a incorporar o cálculo da meia-vida de um conjunto de referências, possibilitando a estimativa da obsolescência do conhecimento em diferentes domínios temáticos. Esse indicador, quando articulado à Análise de Domínio, oferece uma visão mais crítica não apenas da longevidade dos referenciais teóricos, mas também dos fatores institucionais e epistemológicos que influenciam sua manutenção ou substituição.

Segundo investigações clássicas em bibliometria, a meia-vida da literatura em Ciência da Informação (CI) geral situa-se em cerca de 12 anos. Entretanto, estudos de recorte temático revelam variações significativas. Maroldi, Lima e Hayashi (2018) estimaram uma meia-vida de 18,83 anos para publicações na área de Educação, sugerindo maior permanência de seus referenciais. Em contraste, Gabriel (2014), ao analisar citações em um horizonte de dois anos em revistas científicas, identificou que a meia-vida na CI propriamente dita é de apenas seis anos, indicando renovação conceitual mais acelerada nesse campo.

SUMÁRIO

Essas discrepâncias evidenciam que a obsolescência do conhecimento não é homogênea: depende tanto do subcampo temático quanto do intervalo temporal considerado. Ao combinar métricas de meia-vida com a Análise de Domínio, pesquisadores e gestores de acervos podem delinear políticas de fomento mais aderentes às dinâmicas de cada área, orientar a atualização de currículos de pós-graduação e priorizar iniciativas editoriais alinhadas aos ritmos de renovação do saber.

Em estudos de citações, extraem-se diversos indicadores, dentre os quais o Fator de Impacto, proposto por Garfield (1972), é frequentemente apontado como a forma mais simples de aferir a reputação de um periódico. Contudo, uma das maiores críticas ao Fator de Impacto é seu foco na revista, e não na avaliação do pesquisador: publicar em um periódico com FI 5 não significa que o autor detenha índice equivalente. Em resposta a essa limitação, têm ganhado espaço métodos de valoração a nível de autor, capazes de mensurar de forma mais precisa o desempenho individual na produção científica, como o Índice H.

O Efeito Mateus, descrito por Merton (1968), descreve o fenômeno de vantagem cumulativa segundo o qual pesquisadores, instituições ou periódicos já consolidados passam a receber progressivamente mais crédito e visibilidade, reforçando ainda mais sua reputação. Nesse cenário, o Fator de Impacto de um periódico reflete não apenas a qualidade intrínseca dos artigos publicados, mas também esse ciclo de crédito acumulado: quanto maior o FI, maior a probabilidade de novas citações e, conseqüentemente, de elevação contínua desse indicador. Dessa forma, autores que publicam em periódicos de alto FI tendem a ampliar ainda mais sua visibilidade, independentemente de eventuais flutuações na qualidade editorial. Esse viés acumulativo evidencia a necessidade de complementar o Fator de Impacto com métricas capazes de mensurar o mérito individual de autores e artigos, reduzindo a vantagem excessiva conferida às publicações já bem posicionadas.

SUMÁRIO

OUTRAS FERRAMENTAS
PARA PESQUISADOR

Para usuários logados em contas gratuitas, a Brapci oferece diversas ferramentas bibliométricas, como a conversão de listas de autores ou de palavras-chave para o formato .net, visando à construção e análise de redes sociais; ferramentas de contagem de palavras e autores; e o cálculo da vida média de uma lista de referências submetida à análise. Além disso, é possível acessar funcionalidades fundamentadas em cálculo de amostragem e na aplicação da Lei de Price, da Lei de Lotka ou da Lei de Bradford — recursos muito úteis ao analisar grandes volumes de dados.

A proposta da Brapci não é ser apenas mais uma base de dados, mas possibilitar a realização de estudos e a aplicação de metodologias com dados confiáveis, atuando cada vez mais como um laboratório de uso de ferramentas de inteligência artificial em seu contexto de coleta, seleção, organização e acesso à informação especializada na área.

CONSIDERAÇÕES FINAIS

Ao disponibilizar dados bibliográficos enriquecidos e ferramentas analíticas avançadas, a Brapci cumpre papel estratégico na promoção da transparência e da rigidez metodológica em Estudos Métricos da Informação. Sua capacidade de oferecer indicadores diacrônicos, aliados a filtros sincrônicos de disciplinaridade e recortes temporais, permite ao pesquisador não apenas quantificar a produção científica, mas também contextualizar seu significado histórico e temático.

SUMÁRIO

Adicionalmente, a diversidade de funcionalidades, desde a exportação em formatos interoperáveis até a aplicação de leis clássicas da bibliometria (Price, Lotka, Bradford) e de análises multivariadas, transforma a plataforma em um laboratório de experimentação metodológica. Isso favorece a adoção de abordagens complementares que combinam análises quantitativas e qualitativas, mitigando vieses de cobertura e ampliando a visão sobre redes de colaboração e padrões de citação.

Não obstante os avanços, permanece o desafio de assegurar a completude e a regularidade dos registros, especialmente para periódicos com publicações irregulares ou para dados muito recentes, cuja sazonalidade pode distorcer indicadores anuais. A continuidade do desenvolvimento de algoritmos de validação e de técnicas de inteligência artificial para atualização automática de metadados é uma forma de viabilizar a confiabilidade e a abrangência das análises, porém muito trabalho ainda precisa ser feito.

Por fim, ao se consolidar como ambiente de apoio à pesquisa e à gestão científica em português e espanhol, a Brapci contribui para reduzir desigualdades geográficas e sociais na visibilidade acadêmica. Espera-se que sua evolução incorpore inovações tecnológicas e padrões abertos, fomentando redes colaborativas e subsidiando políticas de avaliação e de fomento científico mais inclusivas e contextualizadas. Além disso, ao oferecer uma base teórica e empírica para os Estudos Métricos da Informação, a plataforma enriquece a análise ao contextualizar os dados, indo além da mera quantificação.

REFERÊNCIAS

ARAUJO, C. A. A. Estudo bibliométrico sobre a incidência de dez dos principais autores da Ciência da Informação nos periódicos brasileiros entre 2003 e 2007. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 10., 2009. **Anais** [...] X Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação, 2009.

SUMÁRIO

ARAUJO, R. F. Mídias sociais e comunicação científica: análise altmétrica em artigos de periódicos da ciência da informação. **Em Questão**, v. 21, n. 1, 2015.

ARBOIT, A. E.; BUFREM, L. S. Evolução quantitativa da produção de artigos de periódicos em Ciência da Informação: um estudo cientométrico. **Anais [...]** Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação, v. 11, 2010

BRADFORD, Samuel C. Sources of information on specific subjects. Engineering: **An Illustrated Weekly Journal**, v. 137, p. 85–86, 1934.

BRAGA, G. M. Relações bibliométricas entre a frente de pesquisa (research front) e revisões da literatura: estudo aplicado a ciência da informação. **Ciência da Informação**, v. 2, n. 1, 1973.

BUFREM, L. S.; COSTA, F. D. O.; GABRIEL JUNIOR, R. F.; PINTO, J. S. P. Modelizando práticas para a socialização de informações: a construção de saberes no ensino superior. **Perspectivas em Ciência da Informação**, v. 15, n. 2, 2010.

BUFREM, L. S.; GABRIEL JUNIOR, R. F.; GONCALVES, V. Práticas de co-autoria no processo de comunicação científica na pós-graduação em ciência da informação no brasil. **Informação & Informação**, v. 15, n. 1, 2010.

SILVEIRA, M. A. A.; CAREGNATO, S. E.; BUFREM, L. S. Estudo das razões das citações na ciência da informação: proposta de classificação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 7, n. 2, 2014.

BUSH, V. As we may think. **The Atlantic Monthly**, Julho, 1945.

CAJAZEIRA, P. E. S. L.; SILVA, H. A. As publicações em coautoria e colaboração científica em comunicação na Universidade Federal do Piauí (UFPI). **Ciência da Informação**, v. 50, n. prelo, 2021.

FERREIRA, A. G. C. Bibliometria na avaliação de periódicos científicos. **DataGramaZero**, v. 11, n. 3, 2010.

FIGUEIREDO, L. M. Distribuição da literatura geológica brasileira: estudo bibliométrico. **Ciência da Informação**, v. 2, n. 1, 1973.

FONSECA, E. N. Relações bibliométricas entre a frente de pesquisa (research front) e revisões da literatura: um estudo aplicado à ciência da informação. e, distribuição da literatura geológica brasileira: um estudo bibliométrico. **Revista de Biblioteconomia de Brasília**, v. 1, n. 2, 1973.

SUMÁRIO

FREEMAN, Linton C. Centrality in social networks: conceptual clarification. **Social Networks**, v. 1, n. 3, p. 215–239, 1979.

GABRIEL JUNIOR, R. F. Geração de indicadores de produção e citação científica em revistas de Ciência da Informação: estudo aplicado à base de dados Brapci. 2014. 140 f. **Tese** (doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, 2014.

GABRIEL JUNIOR, R. F. Conjunto de dados com todos os trabalhos da Brapci. Deposita Dados Ibict, 2025. DOI:10.48472/deposita/YVDAKQ

GARFIELD, Eugene. Citation indexes for science; a new dimension in documentation through association of ideas. **Science**, v. 122, n. 3159, p. 108–111, Jul. 1965.

GLÄNZEL, W., MOED, F. H. Journal impact measures in bibliometric research. *Scientometrics*, Netherlands, v. 53, n. 2, p. 171–193, 2002.

GLÄNZEL, Wolfgang. **Bibliometrics as a Research Field: A Course on Theory and Application of Bibliometric Indicators**. [S.l.]: KU Leuven, Jan. 2003. Apostila de curso.

GOUVEIA, F. C. Almetria: métricas de produção científica para além das citações. **Linc** em revista, v. 9, n. 1, 2013.

HIRSCH, Jorge E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 46, p. 16569–16572, Nov. 2005.

HJØRLAND, Birger; ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science**, v. 46, n. 6, p. 400–425, jul. 1995.

LOTKA, Alfred J. The frequency distribution of scientific productivity. **Journal of the Washington Academy of Sciences**, v. 16, n. 12, p. 317–323, June 1926.

MAIA, E. L. S. Comportamento bibliométrico da língua portuguesa, como veículo de representação da informação. **Ciência da Informação**, v. 2, n. 2, 1973.

MAROLDI, A. M.; LIMA, L. F. M.; HAYASHI, M. C. P. I. Vida média e obsolescência da literatura em educação indígena. **InCID: Revista de Ciência da Informação e Documentação**, v. 9 n. 1, n. 1, 2018.

SUMÁRIO

MERTON, Robert K. The Matthew effect in science. **Science**, v. 159, n. 3810, p. 56–63, Jan. 1968.

MOED, Henk F. **Citation analysis in research evaluation**. Dordrecht: Springer, 2005.

NEWMAN, M. The structure of scientific collaboration networks. **Proceedings of the national academy of sciences**, [s.l.], v. 98, n. 2, p. 404–409, 2001.

OLIVEIRA, E. F. T. **Estudos métricos da informação no Brasil**: indicadores de produção, colaboração, impacto e visibilidade. Marília: Oficina Universitária; São Paulo (SP): Cultura Acadêmica, 2018.

OLIVEIRA, E. F. T.; ALVES, B. H.; GRACIO, M. C. C. Análise bibliométrica da revista scire no período de 1996/2010. **Scire**: representación y organización del conocimiento, v. 18, n. 2, 2012.

PEARSON, Karl. Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, v. 187, p. 253–318, 1896.

PRICE, Derek J. de Solla. **Little Science, Big Science**. New York: Columbia University Press, 1963.

PRICE, Derek J. de Solla. **Little Science, Big Science**—and Beyond. New York: Columbia University Press, 1986.

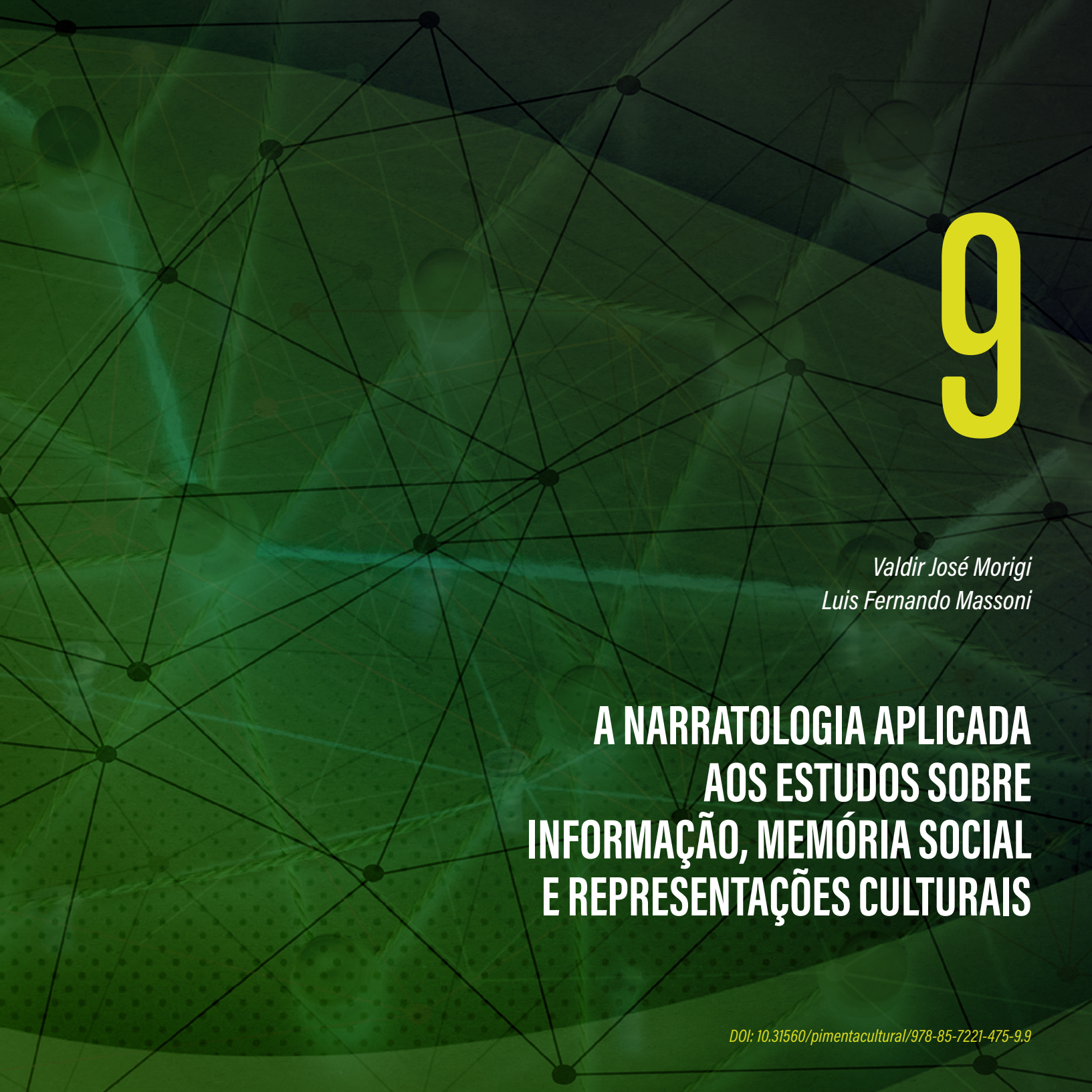
PRITCHARD, Alan M. Statistical Bibliography or Bibliometrics? **Journal of Documentation**, v. 25, n. 4, p. 348–349, dez. 1969.

QUEIROZ, D. G. C.; VILAN FILHO, J. L. A autoria única nas publicações científicas: uma revisão de literatura. **Em Questão**, v. 25, n. 3, 2019.

SMALL, Henry. Co-citation in the scientific literature: a new measure of the relationship between two documents. **Journal of the American Society for Information Science**, v. 24, n. 4, p. 265–269, Jul. 1973.

SPINAK, Ernesto. **Diccionario Enciclopédico de Bibliometría, Cienciometría e Informetría**. Caracas: UNESCO, 1996.

ZIPF, George Kingsley. **Human behavior and the principle of least effort**: an introduction to human ecology. Reading: Addison Wesley, 1949.



9

*Valdir José Morigi
Luís Fernando Massoni*

**A NARRATOLOGIA APLICADA
AOS ESTUDOS SOBRE
INFORMAÇÃO, MEMÓRIA SOCIAL
E REPRESENTAÇÕES CULTURAIS**

DOI: 10.31560/pimentacultural/978-85-7221-475-9.9

SUMÁRIO

COMO CITAR:

MORIGI, Valdir José; MASSONI, Luis Fernando. A narratologia aplicada aos estudos sobre informação, memória social e representações culturais. /n: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

Participamos do livro *Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações*, escrevendo o capítulo sobre as metodologias de pesquisa em memória social e as representações culturais na Ciência da Informação. O uso de metodologias para o estudo da informação, da memória social e das representações culturais, com foco nas práticas informacionais e nos impactos das Tecnologias de Informação e Comunicação (TIC) na vida social, constitui um tema importante e atual para o desenvolvimento da Ciência da Informação.

Acumulamos uma experiência de mais de duas décadas pesquisando temas relacionados à informação, à memória social e às representações culturais, abordando as informações produzidas, publicadas e divulgadas por diferentes atores sociais. Nesse percurso, experienciamos diversas alternativas metodológicas, de acordo com os dados que se apresentavam em cada estudo.

Na condição de uma área do conhecimento relativamente jovem, a Ciência da Informação ainda experimenta a importação de métodos de outras áreas e caminha para desenvolver os seus próprios processos metodológicos. Esse aspecto, que poderia ser visto como fraqueza, na verdade se converte em fortaleza da área:

SUMÁRIO

a liberdade de observar experiências pregressas, analisar os desafios e as oportunidades enfrentadas por cada objeto de estudo e em cada área do conhecimento, de modo a propor abordagens metodológicas que lhe sejam efetivamente válidas.

A memória é um fenômeno complexo e estudá-la por meio das fontes de informação nos apresenta o desafio de compreender os processos de lembrança e esquecimento e as representações que tensionam e reforçam imagens sobre o que chamamos de “memória”. No plano sociocomunicacional, a informação atua na mediação de representações que constroem memórias socialmente partilhadas, o que ocorre pela mídia, mas também por meio de qualquer registro ou discurso oral, escrito ou audiovisual.

Entendendo que tais registros constituem narrativas sobre a experiência humana, importamos do campo das Letras o método da narratologia, convencionalmente empregado na análise de narrativas literárias, mas que pode ser empregado na análise de variadas expressões narrativas. A partir daí, centralizamos este estudo na seguinte questão: como a narratologia pode ser aplicada nos estudos da informação, da memória e das representações culturais? Essa indagação norteia o presente capítulo.

Para tanto, conceituamos o que é a narratologia, a partir da literatura científica sobre o tema, para então apresentarmos algumas possibilidades de aplicação do método nos estudos sobre informação, memória social e representações culturais no âmbito da Ciência da Informação, a partir de uma pesquisa bibliográfica na Base de Dados em Ciência da Informação (BRAPCI). Esperamos que esse texto apresente caminhos e, acima de tudo, provoque indagações.

SUMÁRIO

NARRATOLOGIA COMO MÉTODO

O linguista e filósofo Tzvetan Todorov (1939-) foi quem cunhou o termo francês *narratologie*. O autor concebeu uma transformação de enfoque da narrativa textual de nível superficial (isto é, do nível do discurso concreto construído na forma de letras, palavras e frases) para uma lógica geral e com propriedades estruturais da narrativa, tendo como pano de fundo um universo de representações. Assim, tentou criar uma teoria geral da narrativa com múltiplas possibilidades que pudesse ser aplicada a todos os seus domínios, postulando uma "ciência que ainda não existe; vamos chamá-la de NARRATOLOGIA, ou ciência da narrativa" (Todorov, 1969, p. 10).

Todorov (1939-2017) foi um dos principais expoentes do estruturalismo francês e também um dos críticos desta corrente de pensamento que teve origem na psicologia e que se propagou para a linguística e em diversas ciências sociais, como a antropologia e a sociologia, expandindo para a filosofia do século XX. Segundo Todorov (2006), o livro *As Estruturas das Narrativas* tem como objetivo teórico mostrar como a análise estrutural da narrativa se manifesta no plano abstrato, e não como uma descrição concreta da realidade.

O objetivo de Todorov (2006) é dar o fundamento de uma gramática da narrativa, mas não no sentido de estabelecer normas ou padrões para uma narrativa gramatical, e sim na direção da produção do conhecimento e da classificação das estruturas narrativas. Assim, se poderia desvendar quais as estruturas que existem subjacentes a toda narrativa, estabelecendo uma coletânea de funções, de intrigas e de visões de mundo que a envolvem.

Conforme o autor, "meu objetivo continua sendo sempre não o conhecimento de tal romance ou de tal novela, mas o esclarecimento de um problema geral da literatura ou mesmo dos estudos literários. O estatuto dessa 'teoria literária' não me é, entretanto,

SUMÁRIO

perfeitamente claro" (Todorov, 2006, p. 17). Para ele, a linguagem interfere na construção de gêneros literários, especialmente na teoria literária. A literatura, por si só, é um sistema de signos e na construção do seu discurso literário é possível encontrar muitos traços comuns e similares a todas as linguagens. Assim, o autor explora essa aproximação em *Linguagem e Literatura* e *A Gramática da Narrativa*. Na perspectiva do autor, o texto narrativo possui uma estrutura básica composta por Introdução, desenvolvimento e conclusão; isto é, o texto possui começo, meio e fim. Para ele, a narrativa é constituída das seguintes categorias - ou elementos:

- a. personagens: os sujeitos ou seres que realizam as ações sociais;
- b. enredo: as ações que se desenrolam em um determinado contexto no mundo da narrativa fictícia;
- c. narrador: o centro da narrativa, que pode ser em 1ª pessoa ("eu") ou em 3ª pessoa ("ele/ela");
- d. tempo: a época, período e o modo como a situação ocorre, que pode ser cronológico ou psicológico;
- e. espaço: o lugar (físico ou mental) onde acontece a história;
- f. analepse/prolepse: as projeções no tempo (os retornos ou os avanços);
- g. discurso: as falas dos personagens e os diálogos criados;
- h. ornamentos: são as figuras de linguagem utilizadas para estilizar a tessitura narrativa, tal como uso de metáforas, analogias ou comparações, ironias, entre outros.

Conforme Meister (2022, p. 1), em sua origem, a narratologia teve influência das abordagens estruturalistas, no qual foram desenvolvidas diversas teorias, conceitos e procedimentos metodológicos e analíticos.

SUMÁRIO

Seus conceitos e modelos são amplamente usados como ferramentas heurísticas, e os teoremas narratológicos desempenham um papel central na exploração e modelagem de nossa capacidade de produzir e processar narrativas em uma infinidade de formas, mídias, contextos e práticas comunicativas.

Entretanto, com o seu desenvolvimento do método e de sua teoria ao longo do tempo, houve significativas transformações nas direções e aplicações dessa proposta metodológica. Para tornar o método mais compreensível, convém explicar o caráter central do tempo como elemento articulador da narrativa. Sobre a dimensão temporal (temporalidade), conforme Todorov (2006, p. 21):

A narrativa se constitui na tensão de duas forças. Uma é a mudança, o inexorável curso dos acontecimentos, a interminável narrativa da 'vida' (a história), onde cada instante se apresenta pela primeira e última vez. É o caos que a segunda força tenta organizar; ela procura dar-lhe um sentido, introduzir uma ordem. Essa ordem se traduz pela repetição (ou pela semelhança) dos acontecimentos: o momento presente não é original, mas repete ou anuncia instantes passados e futuros.

Para o autor, a narrativa não obedece a uma ou a outra força, pois se constitui na tensão entre elas. A dimensão temporal da narrativa dirige ao seu caráter episódico, sua fragmentação e ao seu encadeamento. De forma abrangente, a narrativa é composta por um início, meio e fim, abarcando pequenas unidades, os episódios ou incidentes que auxiliam como conectores de uma rede maior.

Foi Müller (1948 *apud* Meister, 2022, p. 9), que introduziu uma distinção fundamental entre "tempo narrado" (*erzählte Zeit*) e "tempo de narração" (*Erzählzeit*), dimensões que caracterizam o ritmo de uma narrativa por meio da sua correlação. Para Meister (2022, p. 9), "essa abordagem foi explorada por Lämmert (1955), em uma das primeiras taxonomias narrativas em grande escala.

SUMÁRIO

Para Lämmert, a fenomenologia das narrativas individuais pode ser rastreada até um repertório universal estável de modos elementares de narrar." Ele fez uma distinção entre os diversos tipos de narração que ora se ampliavam, resumiam, retiravam ou se repetiam, saltavam, paravam, finalizavam, eliminando subsequências, ao passo que outros tipos de narrativa reproduzem plenamente o fluxo do tempo narrado. Essa perspectiva arrolou "essas formas elementares de temporalidade narrativa aos principais modos de narração, como apresentação cênica, relato, reflexão e descrição." (Meister, 2022, p. 9). No entanto, o saldo de sua colaboração teve prejuízo, por ser "uma taxonomia excessivamente complexa e às vezes 'confusa' que tenta explicar todas as formas de flashbacks e flashforwards narrativos." (Meister, 2022, p. 9-10).

Como foi colocado acima, existem muitas taxonomias de entender o tempo na narrativa, por questões de limitações não abordaremos todas elas, mas as mais utilizadas nas análises de narrativa. Sendo o tempo narrativo um dos elementos importante na criação do texto, faremos menção sobre algumas dessas classificações recorrentes nas pesquisas que utilizam a análise da narrativa como procedimento metodológico.

O tempo narrativo pode se dividir em três tipos: tempo cronológico (histórico), tempo imaginário (psicológico) e tempo mítico (metafísico). A concepção do tempo cronológico é definida como o tempo do relógio e calendários. O tempo é dividido em frações ou fragmentos exatos, que vê a passagem do tempo a partir da contagem dos segundos, minutos e horas. Ele está presente na escolha da época em que se transcorre a história, a partir da contagem das datas como: dias, semanas, meses, anos em que as ações dos personagens ocorrem e movimentação e alusão a esses elementos em um texto. Trata-se de um tempo linear que precisa seguir um curso, uma trajetória, tal como se apresentam as horas do relógio. É um tempo que não possui regresso. Isto é, ele não regressa.

SUMÁRIO

O tempo imaginário, denominado também de psicológico, é um tempo significativo que pode alterar a ordem dos acontecimentos, não é um tempo que segue uma ordem linear dos fatos: pode seguir o fluxo do pensamento e se manifestar da evocação de lembranças e imagens sobre o passado. Na perspectiva de Castoriadis (1982), o tempo imaginário é inapreensível e indefinível. Ele não fica à margem do tempo histórico, pode assumir características do tempo identitário (cronológico), mas é uma fenda deste último, uma vez que cada sociedade cria para si um tempo imaginário. Isto é, cada sociedade tem um modo próprio de criar, de lidar e de viver o tempo.

Nessa linha de pensamento, o presente, o passado e o futuro podem se mesclar na narrativa. Isso ocorre porque o tempo imaginário segue o fluxo da consciência ou das lembranças dos personagens, sendo possível analisar as memórias e as representações culturais destes. Ao se evocar o passado por meio de um conjunto de imagens e lembranças sobre um determinado tempo de alguém ou lugar que vem as memórias, se revelam momentos e acontecimentos que ocorreram no tempo cronológico ou identitário. Essa maleabilidade do tempo imaginário, de usufruir as características do tempo identitário, possibilita que este apareça não na forma de uma simples repetição dos fatos, mas como uma manifestação de uma determinada ordem simbólica instituída pela sociedade e das forças que a animam, sejam estes os momentos significativos ou privilegiados das atividades sociais, quer seja na sociabilidade do trabalho, da política, das comemorações, celebrações festivas, entre outros rituais culturais (Castoriadis, 1982).

O tempo mítico, também denominado metafísico, é um tempo coletivo, um tempo cíclico que se manifesta através da natureza e guia, orientando a vida das pessoas. Ele se caracteriza por ser um espelho da natureza arcaica, dos primórdios, sendo reversível por intermédio do retorno ritual às origens. Nele também não há uma divisão entre Deus, natureza e homem, possuindo uma estrutura circular, projetado a partir dos elementos e movimentos cíclicos da

SUMÁRIO

natureza tal como as estações do ano, fases lunares, o plantio e a colheita de alimentos. É considerado um tempo fantasioso, pois não se fundamenta nos padrões do mundo real. Os mitos, os arquétipos, as lendas e o folclore de diferentes culturas são elementos que auxiliam na criação de narrativas que podem ser estudadas como forma de entender como se expressa o tempo mítico. Ele também pode se entrecruzar com o tempo cronológico ou histórico. A concepção do tempo mítico explica o mundo por meio das representações simbólicas e imaginárias, diferente do raciocínio e das experimentações científicas.

Conforme afirma Gomes (2009, p. 59), “[...] o mito se refere a um tempo primordial, tempo que não tem a ver com a história real, mas com um tempo atemporal e originário e que, por isso, representa a totalidade do tempo [...]. Além disso, o mito “é um arquétipo de ação, de cada ação importante, primordial, à qual o homem se refere e que imita nas suas atitudes quotidianas. Se o mito significa o que diz, o seu carácter narrativo supõe um dinamismo que vai do princípio da falta até ao fim da culpabilidade, da perdição do homem à sua salvação.”

O uso cada vez mais intensivo de dispositivos tecnológicos digitais têm provocado transformações nas nossas representações sobre o mundo e com os outros, e consequentemente, afeta a nossa percepção em relação à temporalidade. As TIC fazem com que o ritmo do tempo se transforme, passando a impressão de que o tempo passa de forma acelerada. O tempo linear, intermediado pela duração, passa a ser um tempo atomizado em que os espaços de duração desaparecem, liberando os fatos e os acontecimentos que ficam dispersos, sem experimentar a duração. O que vale é a novidade, fortalecida por um fluxo de informação constante, o que leva a uma desatualização cada vez mais rápida da informação, fazendo com que o tempo perca o seu aroma (Han, 2016).

SUMÁRIO

As múltiplas teias das conectividades das tecnologias digitais no ciberespaço ajudam os espaços culturais (arquivos, bibliotecas, museus entre outros) na manutenção das conexões, tornando as redes sociais espaços complementares para divulgar os seus acervos. Nessa perspectiva, as tecnologias digitais são ferramentas que possibilitam uma nova forma de comunicação, ao mesmo tempo que auxiliam no processo de democratização e do acesso às informações, permitindo novas interações entre os sujeitos com o tempo, com o espaço e com os objetos culturais.

Entretanto, na perspectiva de Han (2016), vivemos em uma era “pós-narrativa”, isto é, caracterizada pelo imediatismo da informação, na qual a internet e as redes sociais são instrumentos de exploração e manipulação. A utilização das redes sociais na sociedade capitalista contemporânea tem como objetivo estimular a competição entre as pessoas e, conforme o autor (2019), as mídias digitais fazem com que as imagens pareçam mais bonitas e melhores do que a própria realidade, o que pode ser uma reação de proteção ou fuga. Assim, sob efeito das TIC, as narrativas sobre os patrimônios culturais digitais são carregadas de informações; ao mesmo tempo que informam, também podem trazer um cansaço de informação.

As narrativas digitais sobre os fenômenos culturais acerca dos patrimônios culturais estão repletas de conteúdos significativos relacionados aos valores e aos sentimentos coletivos, oferecendo tais conteúdos para constituição da cultura, da memória e dos imaginários sociais. No entanto, segundo Han (2016/2019), vivemos atualmente em uma “crise da narração”, era pós-narrativa, auxiliada pelo processo de digitalização, onde praticamente tudo nos é revelado, não havendo mais fantasia narrativa, pois escondemos o vazio de sentido da vida, compartilhando, curtindo ou publicando mensagens nas redes sociais, que deixaram de ser narrativas, mas “informações visuais que desaparecem rapidamente” (Han, 2023, p. 25). Diferentemente de outros tempos, em que as narrativas davam sentido, alicerce que guiavam a vida, completamente oposto de hoje com as micronarrativas fragmentárias sobre o mundo presente.

SUMÁRIO

Como vimos, a narrativa no decorrer de sua história passou por profundas transformações e, com elas, houve uma necessidade de mudar o ritmo do tempo, o qual foi recortado em unidades descontínuas. O tempo linear, que intermediava a duração do tempo, passa ser um tempo atomizado, pois os espaços de duração desaparecem, liberando os fatos e acontecimentos, que se dispersam sem experimentar a duração, tornando-se um tempo sem aroma. Assim se caracterizam as narrativas digitais sobre os patrimônios culturais no mundo contemporâneo.

Para mostrar como ocorre esse fenômeno na Ciência da Informação, trazemos alguns estudos realizados por nós como exemplos da metodologia de análise de narrativas aplicadas nos estudos da informação, da memória social e das representações culturais.

APLICAÇÃO DA NARRATOLOGIA NOS ESTUDOS DA INFORMAÇÃO, DA MEMÓRIA SOCIAL E DAS REPRESENTAÇÕES CULTURAIS

A origem e o uso habitual da narratologia teve início nos estudos literários, mas as suas aplicações são as mais diversas, pois as narrativas são múltiplas e estão em muitos lugares, não apenas em suportes materiais tradicionais, como livros, jornais, revistas etc., pois também se expressam nas variadas dimensões das nossas vidas. A narratologia não se atém somente à análise de narrativas literárias: para Reis (2006), desde o início na década de 1960, o propósito do método incluía a análise de diversos campos, como o mito, o cinema, a publicidade, a narrativa policial, dentre outros. Com o processo de interdisciplinarização dos estudos narrativos, percebemos uma requalificação do método, que passou a ser aplicado em diversos contextos, engendrado nos estudos sobre narratologias cognitivas, narrativas em linguagem natural, feministas, de gênero, midiáticas etc.

SUMÁRIO

Conforme Motta (2013), o objetivo da narratologia é compreender como construímos de modo intersubjetivo os nossos significados via apreensão das representações sobre a realidade, manifestando-se como um método de análise de práticas socioculturais. Para o autor, as narrativas devem ser estudadas a partir das nossas representações sobre o mundo e as coisas (percepções) para compreendermos quem somos e identificarmos como discursos narrativos representam o real e o irreal, seja nas narrativas literárias, midiáticas ou historiográficas. Desse modo, é possível compreendermos de que maneira os sujeitos, os grupos e a sociedade agem diante dos discursos narrativos massivos, consensuais ou controversos discordantes.

As narrativas contam uma história e se constituem por um suceder de fatos, acontecimentos que ocorrem em um tempo definido, com personagens definidos, em um determinado lugar e inseridas em um contexto ou pano de fundo. Elas são formadas por elementos que, quando articulados, são responsáveis pela coerência da história perante o sentido desejado pelo narrador. Para a aplicação da narratologia, recorremos a Motta (2013), para quem a análise empírica das narrativas requer alguns procedimentos:

- a. compreender a *intriga*: diz respeito à análise de como o *enredo* organiza as partes do texto, os seus encadeamentos básicos e os seus contornos da narrativa. Envolve identificar os recursos de linguagem utilizados na narrativa (hipérboles, metáforas, comparações e ironias). Engloba considerar recursos visuais, tais como mapas, gráficos, fotos etc. Além disso, é preciso identificar quais conflitos dramáticos de teor político, ideológico, psicológico, religioso perpassam a narrativa, verificando como o narrador se posiciona diante do tema abordado e os enquadramentos por ele realizados;
- b. entender a *lógica da narração*: perceber de que modo as partes se articulam na narrativa em seu contexto comunicativo, como um projeto dramático constituinte de uma realidade.

SUMÁRIO

Analisar ações, tensões, surpresas, clímax, o começo, o meio e o fim - nem sempre as narrativas apresentem esta lógica, pois podem romper com a linearidade;

- c. identificar os *episódios*: eles são unidades temáticas interdiárias, coesas e que relatam ações relativamente autônomas, englobando transformações e progressões no transcorrer da história, sempre conectadas ao todo;
- d. perceber o *conflito dramático*: neles são realizados os enquadramentos, o ponto de vista pelo qual o narrador narra. Os conflitos (psicológicos, interpessoais, políticos, ideológicos) revelam interesses contraditórios e rompimentos de equilíbrio ou estabilidade, gerando tensões;
- e. analisar as *personagens*: são elas que realizam os enfrentamentos, assumem um tipo, traços singulares que caracterizam certos sujeitos arquétipos, às vezes fortalecendo estereótipos. É fundamental compreender porque uma personagem possui qualidades e defeitos e porque de suas atitudes;
- f. identificar as *estratégias argumentativas*: a narrativa nunca é ingênua e neutra, pois é fundamentada na argumentação e movida por propósitos. É preciso perceber o uso intencional de recursos linguísticos, de modo a identificar marcas e pistas que evidenciem a estratégia do narrador e os jogos de poder;
- g. identificar as *metanarrativas*: a narrativa é circunscrita em um fundo ético e moral, atrelado a questões culturais anteriores à narrativa e que a influenciam. As narrativas revelam mitos que habitam nossas metanarrativas culturais (o crime não compensa, a nação é soberana, o trabalho enobrece etc.). Trata-se do pano de fundo em que se desenvolve o enredo sobre determinado assunto.

Conforme Gancho (2002), que estudou esses elementos a partir das narrativas literárias, a estrutura que compõe as narrativas

SUMÁRIO

literárias é composta por cinco elementos essenciais. Há um *enredo*, com um conjunto de fatos narrados na história, também denominado de ação, intriga, tema etc. O enredo é vivenciado por *personagens*, responsáveis por fazer as ações descritas – pode-se afirmar que eles dão “vida” à narrativa. Eles podem ser seres humanos ou outros animais, além de coisas/objetos, que são definidos pelos seus atos ou por aquilo que dizem, apresentando características físicas, sociais, psicológicas, ideológicas e morais. Os personagens atuam em um determinado *espaço*, que é o cenário onde ocorrem as ações, estabelecendo interações com as personagens. Isso ocorre em um determinado *tempo*, o que engloba a época em que a história ocorre, sua duração, envolvendo as relações entre passado, presente e futuro e as suas percepções, o que engloba tanto a cronologia como o tempo psicológico. Por fim, tudo é narrado por um *narrador*, que é o elemento estruturador da história, seja em primeira ou terceira pessoa.

Além desses elementos citados pela autora, também há o *tema*, que expressa a ideia sobre a qual a história é desenvolvida (Gancho, 2002). Na prática, a aplicação dessas categorias deve respeitar as necessidades de cada narrativa que se pretende estudar, necessitando de adaptações para que o método melhor considere as especificidades de cada contexto estudado, pois o método deve se adequar à realidade dos dados pesquisados, e não o contrário.

Na teoria de aplicação da narratologia, Motta (2013) apresenta algumas instâncias de análise pragmática de narrativas, que orientam a análise de aspectos como os recursos de linguagem utilizados (metáforas, comparações, ironias, exclamações etc.) e as metanarrativas que permeiam as histórias narradas (panos de fundo dos discursos). A observação desses elementos possibilita um olhar mais aprimorado a respeito das possíveis intencionalidades por trás da produção das narrativas.

Na perspectiva de Motta (2013), ao utilizarmos a narratologia como um método de análise devemos ter cuidado para não tornar

SUMÁRIO

a análise rígida, seguindo os cânones de proposição formal, consolidados da literatura, muito menos ter receio de inovar e criar novas metodologias. Desta forma, a narratologia pode auxiliar na fundamentação para o desenho metodológico da pesquisa sobre os múltiplos campos do conhecimento e seus fenômenos, dentre os quais se inserem as relações que envolvem a informação, a memória, as representações culturais e a narrativa.

Com objetivo de mostrar as possibilidades da aplicação da narratologia nos estudos sobre informação, memória social e representações culturais no âmbito da Ciência da Informação, foi realizada uma pesquisa bibliográfica na Base de Dados em Ciência da Informação (BRAPCI). A coleta dos dados foi realizada em outubro de 2024 e consistiu em uma busca pelo termo “narratologia” sem limitação temporal e utilizando todos os campos de busca (título, palavras-chave, resumo, texto completo). Como resultado, obteve-se 15 resultados, publicados entre 2006 e 2023. Destes, cinco foram descartados por não serem estudos do campo da Ciência da Informação, mas de áreas como Letras e Comunicação. Optamos por manter os demais resultados, em um total de 10, mesmo que o foco de alguns não fosse especificamente a relação entre informação e memória, como foi possível observar após a leitura de seus resumos.

O *corpus* de pesquisa é formado por dez textos, quais sejam seis artigos de periódicos e quatro trabalhos publicados nos anais do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB). Procedemos à análise dos dados, que envolveu a leitura na íntegra dos documentos, objetivando identificar as diferentes formas de aplicação do método nos diversos contextos de estudo. Por meio da observação dessas nuances, foi possível apontar possibilidades de aplicação da narratologia nos estudos sobre informação e memória.

Nos textos *A prática profissional do bibliotecário em animês* (Morigi; Massoni; Kussler, 2017a) e *Bibliotecários em animês: representações ficcionais e realidade* (Morigi; Kussler; Massoni, 2017b), são

SUMÁRIO

analisados os temas, os enredos, os ambientes, os tempos e os personagens das narrativas de animês. O foco recai sobre os personagens (bibliotecários), com a observação de características físicas, psicológicas e sociais. Essa centralidade da categoria personagens se dá pelo fato do tema do estudo ser representações sobre bibliotecários.

No texto *As narrativas da cidade no aplicativo Porto Alegre Guide*, Massoni, Morigi, Dodebei e Damin (2017) utilizam a narratologia para observar as categorias enredo, episódios, cenários, personagens e sequências cronológicas, sem fazer distinção entre tema e enredo e analisando o primeiro como uma parte do segundo. O foco foi a análise da narrativa sobre a cidade disponibilizada em um aplicativo elaborado pela Prefeitura de Porto Alegre.

Contrapondo a narrativa oficial acima analisada, a dissertação de Luis Fernando Massoni, orientada por Valdir Morigi, deu origem ao texto *A cidade na palma da mão: informações e memórias no aplicativo Foursquare* (Massoni; Morigi, 2017), em que os autores observam as narrativas sobre a cidade construídas pelos cidadãos que se manifestam por meio do aplicativo de geolocalização *Foursquare*. Nele, os próprios usuários publicam dicas sobre a cidade, sem um filtro institucional.

Os dados da dissertação de Luis Fernando também foram analisados no artigo *Cidade informacional e identidade: o porto alegre narrado no Foursquare* (Massoni; Morigi, 2018), onde o enfoque foi a categoria personagens, para análise da identidade do porto alegre a partir do aplicativo. O texto destaca as relações entre narrativas e identidade, aplicando a narratologia para perceber representações, enquadramentos e estereótipos. Os demais elementos das narrativas também são analisados, mas com menos profundidade.

No artigo *Memória, cidade e narrativa musical: as lembranças do “meu pequeno cachoeiro”*, Morigi (2020) aplica o método da narratologia para estudar uma antiga canção do compositor gaúcho Teixeira, onde ele narra paisagens da sua infância. O estudo aplica o método na caracterização do cenário, dos episódios, das sequências temporais, dos enredos e dos personagens.

SUMÁRIO

Já Ramos e Dumont (2020), no texto *O cotidiano dos super-heróis e protagonistas de histórias em quadrinhos em tempos de pandemia: representação social entre o mundo da ficção e o mundo real*, analisam as ações, posturas e decisões dos personagens no que se refere aos cuidados, à prevenção e à manutenção da saúde durante a pandemia de COVID-19. Nesse estudo, os autores analisam as categorias enredo, personagem, tempo, espaço e narrador, observando tanto os textos como as ilustrações.

A tese de Luis Fernando Massoni, orientada por Valdir Morigi, rendeu o texto *Informação e memória nas narrativas turísticas sobre Porto Alegre* (Massoni; Morigi, 2021), onde os autores utilizam os pressupostos da narratologia para analisar narrativas oficiais e não-oficiais sobre a cidade a partir das informações divulgadas pelo setor turístico. Para tanto, analisaram o aplicativo Foursquare (narrativa não-oficial) e o áudio divulgado pela Prefeitura da cidade no passeio de ônibus da Linha Turismo (narrativa oficial). As categorias analisadas foram narrador, tema, enredo, cenários, personagens e tempos. Os autores concluem que:

A narrativa oficial e institucional sobre o turismo prioriza e divulga informações apoiadas na história da cidade, utilizando fontes de informação oficiais, citação de datas e de personagens importantes. Nela, o presente da cidade é resultado das ações do passado, tratando-se de uma memória oficial sobre a cidade.

Na narrativa não-oficial, as informações decorrem de fontes informais, como as impressões captadas facilmente no dia a dia pelo senso comum, possuindo ampla difusão no imaginário social. Embora apresente certo detalhamento em relação aos lugares e às vivências cotidianas dos cidadãos, ela é superficial do ponto de vista histórico (Massoni; Morigi, 2021, p. 13-14).

Por sua vez, o texto *Primeiras representações da identidade de gênero e da diversidade nas histórias em quadrinhos de super-heróis estadunidenses*, de Bari e Ramos (2023), articula a Teoria das

SUMÁRIO

Representações Sociais (Serge Moscovici) às teorias da Ciência da Informação e à Narratologia de Todorov para compreender a inserção de temas relacionados à identidade de gênero nas histórias em quadrinhos. Entretanto, o estudo não utiliza as categorias propostas pelo método da narratologia.

Por fim, no texto *Notas sobre o uso da narratologia nos estudos sobre informação e memória*, Massoni e Morigi (2023) realizam um levantamento de produções do campo da Ciência da Informação que utilizam o método da narratologia, especialmente nos estudos sobre informação e memória. Na visão dos autores, o método se mostrou “válido nos estudos sobre informação e memória ao possibilitar observarmos de que modo as histórias são narradas na tessitura do presente, por meio de informações que rememoram acontecimentos, pessoas e lugares” (Massoni; Morigi, 2023, p. 11). Este texto, apresentado no ENANCIB 2023, após revisão, ampliação e atualização, deu origem ao presente capítulo.

Observamos que diferentes estudos se apropriam de formas variadas da narratologia para análise dos dados de pesquisa. Todos adotam os mesmos elementos da narrativa (narrador, episódios, cenários, personagens e tempos). No entanto, é preciso adequar essas categorias de acordo com cada contexto situacional da pesquisa ou utilizar um termo variante para se referir ao mesmo elemento: o tempo é também denominado de sequência cronológica ou temporal; o espaço também é denominado de cenário ou ambiente; o narrador é analisado como categoria em alguns estudos e em outros não; na maioria das pesquisas, o tema é analisado como um elemento constituinte da categoria enredo. Estas particularidades evidenciam que não há um único padrão de aplicação do método: há um leque de possibilidades de apropriação dos pressupostos da narratologia.

Embora o método possua uma estrutura formal pré-estabelecida de elementos a serem analisados, é importante destacar que o uso da narratologia não precisa se limitar a esses elementos e muito

SUMÁRIO

menos forçar os dados a darem conta de preencher todos eles. Isso porque é no cotejo dos dados, identificando como eles aparecem em um determinado contexto de pesquisa que verificamos a sua necessidade. Assim, se está em análise uma narrativa que possui poucas informações temporais ou não indica referências a lugares ou é difícil caracterizar quem é o narrador, isso não se constitui num impeditivo para a aplicação do método. Observamos que, dentre os trabalhos analisados, alguns utilizaram a narratologia especificamente para analisar personagens, ainda que também tenham feito apontamentos genéricos sobre os demais elementos da narrativa.

Na formulação do problema de pesquisa, por exemplo, o(a) pesquisador(a) pode se ater às categorias utilizadas pela narratologia. Assim, se a pesquisa envolve as representações culturais ou grupo de pessoas, são formuladas questões do tipo: quem são os personagens responsáveis pelas produções das representações? Quais os temas, os enredos, os ambientes? Em que tempo ocorre a narrativa (passado/presente/futuro)? Da mesma forma, a organização dos dados deve se ater às categorias utilizadas pelo método – os temas, os enredos, os ambientes, os tempos e os personagens, como foram apontados anteriormente no texto.

Na prática, esses elementos (narrador, episódios, cenários, personagens e tempos) auxiliam o pesquisador na criação de categorias que podem ser analisadas em uma pesquisa, principalmente na forma como os dados podem ser organizados e apresentados. A partir da criação de categorias analíticas, é possível fazer um quadro sintético de forma que os dados apareçam objetivamente, ou então por meio de um conjunto de parágrafos para cada uma das categorias, auxiliando nas divisões do texto em subseções.

Além disso, as categorias ajudam a dinamizar o fluxo informacional que rege o texto, formam uma linha de raciocínio ou de pensamento, sem desconsiderar os demais recursos narrativos que podem auxiliar esse processo, tais como ilustrações, mapas, fotos, gráficos etc.,

SUMÁRIO

pois eles também possuem informações que podem ser úteis no momento da descrição ou da análise dos dados. A aplicação da narrativa nos estudos sobre informação, memória e representações culturais auxilia na compilação e na organização dos dados da pesquisa.

Entretanto, é fundamental que o(a) pesquisador(a) se atenha ao fato de que os elementos que compõem a narrativa precisam ser analisados em conjunto, e não isoladamente. Por mais que a narrativa proponha a observação da estrutura da narrativa, precisamos lembrar que a narrativa contada é o cerne da questão, sendo moldada de acordo como são articulados os seus elementos. Ou seja, apenas identificar e caracterizar uma personagem não é o suficiente: precisamos compreender qual seu papel nas intrigas, sua interação com o espaço e sua participação no transcorrer do tempo.

Para ilustrar este estudo, o quadro 1 apresenta as categorias adotadas em cada texto recuperado na pesquisa bibliográfica empregada neste estudo.

Quadro 1 – Síntese das categorias analisadas nos estudos analisados

Referência	Categorias analisadas
Bari; Ramos, 2023	-
Massoni; Morigi, 2017	Tema, cenários, personagens, enredos e tempos.
Massoni; Morigi, 2018	Foco na categoria personagens.
Massoni; Morigi, 2021	Narrador, tema, enredo, cenários, personagens e tempos.
Massoni; Morigi, 2023	-
Massoni; Morigi; Dodebei; Damin, 2017	Enredo, episódios, cenários, personagens e sequências cronológicas.
Morigi, 2020	Cenário, episódios, sequências temporais, enredos e personagens.
Morigi; Kussler; Massoni, 2017a	Tema, enredos, ambientes, tempos e personagens.
Morigi; Kussler; Massoni, 2017b	Tema, enredos, ambientes, tempos e personagens.
Ramos; Dumont, 2020	Enredo, personagens, tempos, espaços e narrador.

Fonte: dos autores, 2024.

SUMÁRIO

Os elementos da narratologia evidenciam o que deve ser lembrado na narrativa, mas a memória também é constituída dos esquecimentos. Nesse caso, é preciso que o pesquisador esteja atento para realizar constantes tensionamentos, questionando sobre o que e o porquê algo foi esquecido, procurando identificar quais as causas desses esquecimentos.

O uso do método possibilita verificar a transversalidade dos dados, considerando que existe uma conexão entre os grupos sociais estudados e as escolhas dos suportes em que se encontram as informações (oral, impressa ou eletrônica) ou mesmo do gênero documental (científico, jornalístico, publicitário etc.) e os recursos visuais afetam o sentido de um determinado fenômeno informacional. Isso trará especificidades na análise, permitindo compreender de que modo aparecem os sentidos explícitos ou implícitos das ideologias que compõem o tecido informacional.

É importante destacar que a narratologia é empregada para o estudo de narrativas reais e ficcionais, podendo ser adaptada e utilizada em um amplo espectro de estudos. Entretanto, é fundamental que o(a) pesquisador(a) identifique que os seus dados constituem uma narrativa. Isso porque, os dados em si, meramente coletados e organizados, não formam uma narrativa: temos uma narrativa quando existe um narrador (ou vários) que contam/registram uma história. Entendemos ser este o caso de muitos estudos que envolvem pessoas, grupos ou instituições que produzem informação, memória e representações culturais. O método pode ser aplicado em diferentes contextos sociais, o seu uso na Ciência da Informação pode ser mais eficaz em ambientes sociais em que se produz a informação social.

Além disso, é preciso deixar claro que o uso e a aplicação da narratologia deve ser entendido como um guia que oriente o(a) pesquisador(a) no processo da pesquisa e, portanto, não precisa ser seguido como um padrão metodológico rígido, pois ele é aberto à sensibilidade, à criatividade e às novas possibilidades ou elementos que podem surgir no decorrer da pesquisa.

SUMÁRIO

As pesquisas que utilizam a narratologia como método na área da Ciência da Informação são recentes e escassas no país. Dos 10 estudos encontrados na Brapci, oito deles são de Valdir Morigi e/ou Luis Fernando Massoni, e dois de Rubem Borges Teixeira Ramos, o que mostra que o uso da metodologia ainda está restrito a apenas dois grupos de pesquisadores.

CONSIDERAÇÕES FINAIS

Os estudos teóricos sobre a narrativa iniciaram com o linguista russo Tzvetan Todorov, que, inicialmente, estava interessado em conhecer as estruturas das narrativas literárias, base que deu o alicerce fundador do método. Entretanto, com o passar do tempo, foi apropriado e aplicado em outras áreas do conhecimento.

Os elementos das narrativas (narrador, episódios, cenários, personagens e tempos) estruturam as histórias reais ou ficcionais e são a tessitura que forma o tecido informacional. Os elementos da narrativa permitem que essas informações textuais sejam organizadas e melhor analisadas. Assim, é possível fazer uma leitura da informação como um fenômeno social constituído por narrativas.

Nos estudos sobre memória social, a categoria tempo é fundamental para compreendermos a construção da memória. Assim, é importante identificar qual a concepção de tempo é utilizada nas narrativas estudadas: como vimos, os tempos podem ser classificados em tempo cronológico (histórico), tempo imaginário (psicológico) e tempo mítico (metafísico). Essa identificação pode auxiliar no refinamento da análise, tornando os estudos mais profundos.

O estudo das narrativas no campo da Ciência da Informação, especialmente nos estudos sobre memória e representações culturais, possibilita identificar os enquadramentos das memórias a partir

SUMÁRIO

da mediação de informações, que vêm permeadas pelas âncoras das representações acerca dos fenômenos sociais. As nossas representações sobre a realidade e os fenômenos que nos cercam auxiliam na constituição das nossas visões de mundo, que por sua vez afetam as memórias individuais e coletivas.

A análise de narrativas nos permite desvendar como operam os processos de representação e de constituição da realidade social e histórica, o que implica considerar as circunstâncias situacionais dos atos de fala, a identidade cultural e a posição dos interlocutores no contexto social, bem como suas intencionalidades, estratégias de linguagens argumentativas, discursivas e narrativas e o entorno cultural compartilhado. O estudo de narrativas revela percepções de mundo que estão permeadas de representações culturais, presentes na memória individual e que se interligam com a memória dos grupos sociais.

O emprego da narratologia na Ciência da Informação se aproxima da perspectiva sociocultural sobre o campo, orientando a análise dos dados de modo a pensar os usos sociais, políticos, econômicos, artísticos e culturais da informação. Uma narrativa é uma informação carregada de sentidos que operam nos campos ideológicos, evidenciando práticas informacionais enredadas em disputas e tensões sociais.

Desejamos, com este estudo, que o método seja mais difundido e aplicado nos estudos em Ciência da Informação. Nesse sentido, acreditamos que uma de suas principais potencialidades é a liberdade de abordagens que permite: suas categorias são pré-prontas, mas podem ser organizadas e analisadas de acordo com cada contexto de estudo. Como limitações, apontamos a ausência de produções na área que empreguem o método, o que dificulta a compreensão de modelos de análise aos novos pesquisadores e evidencia um caminho pouco conhecido a ser explorado.

Os estudos sobre informação, memória social e representações culturais são transdisciplinares e complexos, o que justifica a necessidade de se recorrer a múltiplas teorias e propostas metodológicas que

SUMÁRIO

deem conta de explicá-los. Esperamos que este trabalho sirva para dar visibilidade ao método no campo da Ciência da Informação, tendo em vista o incipiente uso identificado no levantamento realizado.

REFERÊNCIAS

BARBOSA, Márcio Ferreira. **Experiência e narrativa**. Salvador: Edufba, 2003.

BARI, Valéria Aparecida; RAMOS, Rubem Borges Teixeira. Primeiras representações da identidade de gênero e da diversidade nas histórias em quadrinhos de super-heróis estadunidenses. **Liinc em Revista**, Rio de Janeiro, v. 19, n. 2, p. 1-24, 2023. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/248490>. Acesso em: 17 out. 2024.

CASTORIADIS, Cornelius. **A instituição imaginária da sociedade**. 2 ed. Rio de Janeiro: Paz e Terra, 1982.

GOMES, Manuel Tavares. Tempo histórico e tempo mítico: a descoberta da identidade no «tempo do sentido». **Percursos**, Florianópolis, v. 10, n. 02, p. 56-77, jul./dez. 2009. Disponível em: <https://revistas.udesc.br/index.php/percursos/article/view/1605/1517>. Acesso em: 08 dez. 2024.

HAN, Byung-Chul. **O aroma do tempo**: um ensaio filosófico sobre a arte da demora. Lisboa: Relógio d'Água. 2016.

HAN, Byung-Chul. **No enxame**: perspectivas do digital. Trad. Lucas Machado Petrópolis, Vozes, 2019.

HAN, Byung-Chul. **A crise da narração**. Petrópolis: Vozes. 2023.

MASSONI, Luis Fernando Herbert; MORIGI, Valdir José. A cidade na palma da mão: informações e memórias no aplicativo foursquare. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017. **Anais [...]**, Marília: UNESP, 2017. p. 1-19. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/104088>. Acesso em: 17 out. 2024.

MASSONI, Luis Fernando Herbert; MORIGI, Valdir José. Cidade informacional e identidade: o porto alegre narrado no Foursquare. **Inf. & Soc.:** Est., João Pessoa, v. 28, n. 1, p. 169-179, jan./abr. 2018. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/91182>. Acesso em: 17 out. 2024.

SUMÁRIO

MASSONI, Luis Fernando Herbert; MORIGI, Valdir José. Informação e memória nas narrativas turísticas sobre porto alegre. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 21., 2021. **Anais [...]**, Rio de Janeiro: IBICT/UF RJ, 2021. p. 1-15. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/193133>. Acesso em: 17 out. 2024.

MASSONI, Luis Fernando Herbert; MORIGI, Valdir José. Notas sobre o uso da narratologia nos estudos sobre informação e memória. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 23., 2023. **Anais [...]**, Aracaju: UFS, 2023. p. 1-12. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/258317>. Acesso em: 17 out. 2024.

MASSONI, Luis Fernando Herbert; MORIGI, Valdir José; DODEBEI, Vera; DAMIN, Marina Leitão. As marrativas da cidade no aplicativo Porto Alegre Guide. **Inf. & Soc.: Est**, João Pessoa, v. 27, n. 1, p. 147-160, jan./abr. 2017. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/91230>. Acesso em: 17 out. 2024.

MEISTER, Jan Christoph. Narratologia. Tradução: Gisele da Costa Lipimam Pereira. **Repositório Digital de Textos de Narratologia**. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2022. Título Original: Narratology. Disponível em: <https://www.ufrgs.br/geni/wp-content/uploads/2023/02/Narratologia.pdf>. Acesso: 01 nov. 2024.

MORIGI, Valdir José. Memória, cidade e narrativa musical: as lembranças do “meu pequeno cachoeiro”. **Inf. & Soc.: Est**, João Pessoa, v. 30, n. 4, p. 1-13, out./dez. 2020. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/156947>. Acesso em: 17 out. 2024.

MORIGI, Valdir José; KUSSLER, Natan Fritscher; MASSONI, Luis Fernando Herbert. Bibliotecários em animês: representações ficcionais e realidade. **Informação & Informação**, Londrina, v. 22, n. 3, p. 320-345, 2017a. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/33750>. Acesso em: 17 out. 2024.

MORIGI, Valdir José; MASSONI, Luis Fernando Herbert; KUSSLER, Natan Fritscher. A prática profissional do bibliotecário em animês. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 18., 2017b. **Anais [...]**, Marília: UNESP, 2017. p. 1-20. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/104301>. Acesso em: 17 out. 2024.

MOTTA, Luiz Gonzaga. **Análise crítica da narrativa**. Brasília, DF: Editora UnB, 2013.

RAMOS, Rubem Borges Teixeira; DUMONT, Lígia Maria Moreira. O cotidiano dos super-heróis e protagonistas de histórias em quadrinhos em tempos de pandemia: representação social entre o mundo da ficção e o mundo real. **Liinc em Revista**, Rio de Janeiro, v. 16, n. 2, p. 1-20, 2020. Disponível em: <https://hdl.handle.net/20.500.11959/brapci/157497>. Acesso em: 17 out. 2024.

SUMÁRIO

REIS, Carlos. Narratologia(s) e teoria da personagem. **Revista do Programa de Pós-Graduação em Letras da Universidade de Passo Fundo**, Passo Fundo, v. 2, n. 1, p. 26-36, jan./jun. 2006. Disponível em: <https://seer.upf.br/index.php/rd/article/view/505/312>. Acesso em: 08 dez. 2024.

SOUZA; Paulo Victor; CUNHA, Rodrigo do Espírito Santo da. Entre o ser e o estar: a representação do eu e do lugar no Foursquare. *In*: RIBEIRO, José Carlos; FALCÃO, Thiago; SILVA, Tarcízio (org.). **Mídias sociais: saberes e representações**. Salvador: EDUFBA, 2012. p. 31-47.

TODOROV, Tzvetan. **Grammaire du Décaméron**. The Hague: Mouton, 1969.

TODOROV, Tzvetan. **As estruturas narrativas**. 4. ed. São Paulo: Perspectiva, 2006.

Eixo

3

**PARADIGMAS
DIGITAIS
E ABERTOS**

A exploração dos paradigmas contemporâneos da era digital orienta os capítulos deste eixo, que abordam metodologias para a prática da Ciência Aberta, a preservação de acervos, a garantia da qualidade dos dados de pesquisa e a aplicação de processos colaborativos como o co-design.

10

Fabiano Couto Corrêa da Silva

METODOLOGIAS PARA A CIÊNCIA ABERTA:

**MÉTRICAS, GOVERNANÇA E PRÁTICAS
NA CIÊNCIA DA INFORMAÇÃO**

DOI: 10.31560/pimentacultural/978-85-7221-475-9.10

SUMÁRIO

COMO CITAR:

SILVA, Fabiano Couto Corrêa da. Metodologias para a Ciência Aberta: métricas, governança e práticas na ciência da informação. In: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações**. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

A Ciência Aberta se destaca como uma frente inovadora e desafiadora no cenário contemporâneo da pesquisa científica, mobilizando múltiplas disciplinas e demandando novas reflexões sobre suas implicações teóricas, éticas e metodológicas. Neste capítulo, exploraremos tanto os fundamentos que sustentam esse movimento quanto os desdobramentos metodológicos, éticos e políticos que emergem de sua implementação em diferentes contextos. A análise proposta enfatiza como essa abordagem redefine os modelos de produção e disseminação do conhecimento, articulando-se em torno de três pontos centrais que exploraremos: as práticas metodológicas, as métricas de avaliação e a governança do ecossistema científico.

Ao defender o acesso aberto e irrestrito ao conhecimento científico sem custos para autores ou leitores, a Ciência Aberta emerge como uma questão central nesse debate. Esse modelo enfatiza a sustentabilidade por meio do apoio institucional, governamental ou comunitário, evitando a dependência de taxas de publicação ou assinaturas. Tal concepção promove, assim, uma democratização efetiva do conhecimento, garantindo que barreiras econômicas não limitem o acesso às informações científicas. Diante desse cenário, torna-se essencial analisar as práticas, impactos e desafios relacionados, especialmente em um contexto global marcado por profundas desigualdades no acesso ao conhecimento.

SUMÁRIO

O estudo desse movimento demanda uma abordagem estruturada que considere a complexidade de suas práticas e conceitos. Ao promover a abertura como princípio fundamental, além de ampliar o acesso a resultados de pesquisas científicas, o movimento também introduz novos paradigmas de colaboração, transparência e inclusão. Sua relevância vai além da adoção de tecnologias ou políticas específicas; trata-se de uma mudança cultural que desafia as normas estabelecidas e incentiva o envolvimento de múltiplos atores, incluindo pesquisadores, cidadãos e formuladores de políticas. Busca-se, portanto, oferecer uma análise abrangente das dimensões que compõem a Ciência Aberta, abordando suas práticas, impactos e desafios. A ênfase recai sobre o potencial transformador dessas práticas para enfrentar problemas globais, como desigualdades de acesso ao conhecimento, reprodutibilidade científica e a conexão da ciência com demandas sociais.

Ao longo do capítulo, estabelece-se um diálogo entre teoria e prática, com a apresentação de técnicas de pesquisa e reflexões sobre como a Ciência Aberta, guiada por princípios de abertura, responsabilidade e colaboração, pode reconfigurar as relações entre produção científica e sociedade. Para isso, o texto discute práticas e impactos, e qualifica desafios contemporâneos como a crise de reprodutibilidade, as desigualdades de acesso e a demanda por retorno social. A partir desse diagnóstico, delimita-se a abordagem metodológica central do capítulo, ancorada na taxonomia latino-americana de Ciência Aberta (Silveira *et al.*, 2023). Essa ferramenta, com suas 10 facetas e 96 rótulos, oferece a base para fundamentar o desenho de estudos, a definição de indicadores (métricas) e as decisões de governança. Estes são os três pontos que estruturam a análise a seguir.

SUMÁRIO

A CIÊNCIA ABERTA COMO *OBJETO DE PESQUISA*

A Ciência Aberta representa um movimento global que promove a transparência, o acesso universal e a colaboração em todas as etapas do ciclo de pesquisa científica. Este caminho não é somente tecnológico e investigativo, mas também ético, regulatório e profundamente político. Trata-se de uma mudança paradigmática que desafia práticas acadêmicas tradicionais, propondo novos modelos para a produção, validação e disseminação do conhecimento. Suas práticas incluem publicação em acesso aberto, compartilhamento de dados, uso de softwares livres e colaboração ativa entre pesquisadores e cidadãos, como em iniciativas de ciência cidadã. Além de democratizar o conhecimento, esse movimento busca reduzir barreiras econômicas e geográficas, tornando a ciência mais acessível e inclusiva.

Embora a popularidade dessa abordagem tenha crescido nas últimas décadas, suas raízes estão ligadas a movimentos históricos de democratização do acesso ao conhecimento. O movimento de acesso aberto, por exemplo, teve marcos importantes nas décadas de 1990 e 2000, com iniciativas como a Budapest Open Access Initiative (2002) e a Berlin Declaration on Open Access (2003) (Morrison, 2006). Tais esforços foram uma resposta às altas taxas de assinaturas de periódicos acadêmicos, que limitavam o acesso à literatura científica, especialmente para pesquisadores de instituições com menos recursos. Desde então, a ideia de abertura na ciência evoluiu para incluir o compartilhamento de dados brutos, metodologias detalhadas e códigos de software, promovendo práticas mais inclusivas e colaborativas no fazer científico.

A importância da Ciência Aberta tornou-se ainda mais evidente durante a pandemia de COVID-19. O compartilhamento rápido

SUMÁRIO

de dados sobre o vírus, incluindo sequenciamentos genéticos e descobertas científicas, foi essencial para o desenvolvimento de vacinas e a implementação de soluções em tempo recorde (Xu *et al.*, 2022). Essa abordagem colaborativa destacou o impacto positivo de uma Ciência Aberta e acessível para a solução de problemas globais, integrando esforços de pesquisadores de diferentes países e permitindo respostas mais ágeis e eficientes.

Apesar de seus benefícios, essa abordagem enfrenta desafios significativos. Há resistência cultural em setores tradicionais da academia, como apontado por Fecher e Friesike (2014), especialmente em relação à abertura de dados e práticas colaborativas que rompem com modelos hierárquicos e competitivos de produção científica. Estudos como o de Pontika *et al.* (2015) também destacam que muitos pesquisadores demonstram incertezas quanto ao reconhecimento acadêmico e à credibilidade de plataformas abertas. Além disso, há a necessidade de capacitação para a adoção de novas práticas e a adaptação de políticas institucionais e governamentais, o que inclui infraestrutura técnica e marcos legais ainda em consolidação. Questões relacionadas à proteção de dados sensíveis, direitos autorais e sustentabilidade financeira também representam obstáculos importantes. No entanto, as perspectivas para o futuro são promissoras, com o crescimento de redes colaborativas, o fortalecimento de políticas públicas em favor da Ciência Aberta e o avanço de tecnologias que facilitam a circulação do conhecimento. Desse modo, o movimento se consolida como eixo estruturante da ciência contemporânea, moldando um ecossistema de pesquisa mais transparente, inclusivo e eficiente, essencial para enfrentar os desafios científicos e sociais da atualidade (Eismann, 2017).

PROBLEMAS CONTEMPORÂNEOS

O movimento da Ciência Aberta se tornou especialmente relevante no cenário atual ao enfrentar problemas que desafiam a ciência global, como a chamada crise de reprodutibilidade, as desigualdades de acesso ao conhecimento e a necessidade de maior retorno social das pesquisas financiadas com recursos públicos.

Essa crise refere-se à dificuldade de confirmar os resultados apresentados em diversos estudos científicos, indicando falhas na transparência e no rigor metodológico. Muitos trabalhos não disponibilizam dados brutos, metodologias detalhadas ou informações suficientes que permitam a verificação independente de suas conclusões (Grácio *et al.*, 2020). Isso compromete a confiança na ciência e impede o avanço cumulativo do conhecimento, pois resultados não confirmáveis dificultam a construção de novas descobertas. Nesse contexto, iniciativas voltadas à abertura da pesquisa científica, ao incentivarem práticas como o compartilhamento de dados e a documentação minuciosa dos métodos, buscam justamente enfrentar esse desafio, promovendo uma cultura de integridade e clareza na produção do conhecimento (Souza; Aganette, 2020).

As desigualdades de acesso ao conhecimento são outro desafio significativo, especialmente em países de baixa e média renda. O custo elevado de assinaturas de periódicos científicos e de acesso a bases de dados cria uma barreira que exclui instituições e pesquisadores de regiões menos favorecidas economicamente. Além disso, essas desigualdades também afetam a capacidade de produção de ciência de qualidade nesses locais, perpetuando ciclos de exclusão e dependência científica (Cabsela, 2023). O movimento da Ciência Aberta combate esse cenário ao defender o acesso aberto às publicações, o uso de softwares livres e o compartilhamento de dados, promovendo uma maior equidade no acesso e na produção do conhecimento.

SUMÁRIO

A demanda por maior retorno social das pesquisas públicas é um reflexo das expectativas crescentes de que a ciência seja mais inclusiva e relevante para a sociedade. Pesquisas financiadas com recursos públicos frequentemente enfrentam críticas sobre sua aplicabilidade prática e sua acessibilidade. A Ciência Aberta responde a essas críticas ao promover práticas que conectam mais diretamente o conhecimento científico com a sociedade, como a ciência cidadã, onde cidadãos participam ativamente de projetos de pesquisa, e o acesso aberto, que garante que os resultados estejam disponíveis para todos. Dessa forma, a ciência se torna não apenas um instrumento de produção acadêmica, mas também uma ferramenta de transformação social.

Como uma resposta estruturada às limitações do sistema científico tradicional, a Ciência Aberta enfrenta questões como confiabilidade, desigualdades e relevância social, redefinindo práticas e reforçando o papel da ciência como um bem público global, capaz de atender às demandas de um mundo interconectado e em constante transformação. Mais do que um ideal, ela se apresenta como uma solução pragmática a esses desafios, promovendo a incorporação de princípios que tornam a ciência mais inclusiva, ética e inovadora, alinhada às necessidades de uma sociedade em rápida evolução.

A RELEVÂNCIA DA CIÊNCIA ABERTA COMO OBJETO DE PESQUISA

Estudar a Ciência Aberta significa investigar as transformações que ela está promovendo no ecossistema científico e suas implicações mais amplas. Além de ser um conjunto de práticas, trata-se também de um campo interdisciplinar e multifacetado, o que a posiciona como um objeto de pesquisa central para compreender

SUMÁRIO

os rumos da ciência na atualidade. Entender como suas iniciativas são implementadas, quais desafios enfrentam e que impactos geram pode contribuir para:

- **Formulação de políticas públicas e institucionais:** Pesquisas podem subsidiar decisões sobre regulamentação de dados, incentivos à publicação aberta e desenvolvimento de infraestruturas científicas.
- **Transformação das práticas científicas:** Estudos ajudam a identificar barreiras culturais e tecnológicas, oferecendo estratégias para promover a adesão de pesquisadores e instituições.
- **Inclusão social e democratização do conhecimento:** A análise de projetos de ciência cidadã e outras iniciativas abertas pode evidenciar seu potencial para engajar comunidades sub-representadas.

Diante dessa relevância e complexidade, torna-se necessário um olhar sistemático sobre os caminhos metodológicos possíveis para investigar a Ciência Aberta no contexto da Ciência da Informação. Para esse fim, será adotada A Taxonomia de Ciência Aberta de Silveira *et al.* (2023) foi elaborada com base na taxonomia do projeto FOSTER (Pontika *et al.*, 2015) e na versão brasileira de 2021, incorporando as Recomendações da UNESCO (2021); foi validada por 21 especialistas de diversos países e resultou em 10 facetas e 96 rótulos.

[illegible]

Estruturada em categorias amplas, a taxonomia apresentada por Silveira, *et al.* (2023) vai além de identificar os componentes-chave da Ciência Aberta, oferecendo também uma análise detalhada das relações entre esses componentes. Essa abordagem permite compreender como cada dimensão interage com as demais, criando sinergias que ajudam a transformar o ecossistema científico. Ao evidenciar essas conexões, a taxonomia destaca o papel estratégico de cada elemento na promoção de práticas mais colaborativas, inclusivas e éticas no cenário da pesquisa contemporânea.

319

SUMÁRIO

Além disso, a taxonomia possibilita identificar desafios e oportunidades para o desenvolvimento de estudos que fortaleçam a Ciência Aberta, oferecendo um panorama completo e estruturado de suas potencialidades.

A seguir, apresentamos as principais dimensões da taxonomia da Ciência Aberta, suas interconexões, exemplos concretos de aplicação e abordagens metodológicas que podem ser empregadas para investigar cada aspecto dessa estrutura dinâmica e interdisciplinar.

ACESSO ABERTO

O Acesso Aberto (AA) é um movimento fundamental para a democratização do conhecimento científico, permitindo que informações e pesquisas sejam acessíveis a todos, sem barreiras financeiras. Esse conceito abrange diferentes vias de acesso, sendo as mais conhecidas a via dourada e a via verde. A via dourada refere-se à publicação de artigos em periódicos de acesso aberto, onde os conteúdos são imediatamente disponíveis para leitura e *download*, enquanto a via verde permite que os autores façam o autoarquivamento de seus trabalhos em repositórios institucionais ou temáticos (Nascimento, 2023). Essa dualidade de modelos é essencial para atender a diferentes necessidades e contextos dentro da comunidade científica. Em nível de abertura mais ampla, tem ganhado destaque o modelo de Acesso Diamante, considerado o expoente mais aberto entre as modalidades, pois garante a gratuidade tanto para leitores quanto para autores, ou seja, não há cobrança de taxas de publicação nem de acesso, sendo geralmente financiado por instituições acadêmicas, consórcios públicos ou políticas de apoio à ciência aberta. Esse modelo fortalece o compromisso com a equidade no ecossistema científico e amplia o alcance do conhecimento, especialmente em países com menor capacidade de investimento individual em publicação científica.

SUMÁRIO

Plataformas como a SciELO (Scientific Electronic Library Online)¹ e o PLOS² (Public Library of Science) têm sido pioneiras na promoção do Acesso Aberto, oferecendo uma ampla gama de periódicos que operam sob esse modelo. A SciELO e o PLOS não apenas garantem que os artigos sejam acessíveis, mas também se comprometem com a transparência e a integridade na publicação científica. Essa prática é reforçada por padrões de dados abertos, que assegurem que os dados associados às pesquisas sejam também acessíveis e reutilizáveis, aumentando a compreensão e a aplicabilidade dos resultados para públicos diversos (Nascimento, 2023).

Os dados abertos têm um papel significativo em várias áreas, especialmente em políticas públicas. Durante a pandemia de COVID-19, a disponibilização de dados em tempo real, como os fornecidos pela Johns Hopkins University (2020), demonstrou a importância do Acesso Aberto para a tomada de decisões informadas e para a resposta a crises de saúde pública. Essa transparência não apenas fortalece a confiança pública, mas também promove a colaboração entre diferentes setores, permitindo que pesquisadores, formuladores de políticas e cidadãos trabalhem juntos em prol de soluções eficazes (Café *et al.*, 2022).

O impacto do Acesso Aberto pode ser analisado por meio de estudos bibliométricos, que avaliam de que forma a disponibilidade gratuita de artigos influencia o número de citações recebidas e estimula colaborações interdisciplinares. Essas análises também permitem investigar a agilidade na comunicação científica, especialmente com o uso de preprints, versões preliminares de artigos disponibilizadas antes da revisão por pares (Muck, 2023). Esse tipo de investigação é fundamental para compreender a eficácia do Acesso Aberto e sua contribuição para um ecossistema de pesquisa mais colaborativo, ágil e acessível.

1 <https://www.scielo.org/>

2 <https://plos.org/>

SUMÁRIO

Além disso, a análise das taxas de processamento de artigos (APCs) em diferentes países revela disparidades no acesso a modelos abertos. Embora o Acesso Aberto tenha crescido significativamente, a cobrança de APCs por alguns periódicos pode limitar a participação de pesquisadores de instituições com menos recursos financeiros (Santos; Monteiro, 2013). Portanto, é fundamental explorar alternativas sustentáveis, como a criação de consórcios institucionais, capazes de mitigar custos e assegurar que o Acesso Aberto seja efetivamente disponível a todos os pesquisadores, independentemente de sua situação financeira.

A partir dessas considerações, é possível identificar oportunidades relevantes para ampliar as investigações sobre o tema. Sugere-se, por exemplo, que estudos futuros analisem modelos alternativos para garantir a sustentabilidade financeira das publicações em Acesso Aberto, especialmente em contextos econômicos menos favorecidos, além de realizar avaliações comparativas sobre como diferentes políticas institucionais influenciam a adesão dos pesquisadores ao Acesso Aberto.

DADOS ABERTOS

Os dados abertos, quando orientados pelos princípios FAIR (Findable, Accessible, Interoperable, Reusable)³, fortalecem a transparência e promovem uma cultura de compartilhamento. Ao viabilizar o reuso, a interoperabilidade e a avaliação independente, esses princípios impulsionam o avanço do conhecimento científico (Grácio *et al.*, 2020).

SUMÁRIO

Iniciativas como o Zenodo⁴ e Figshare⁵ destacam o papel dos repositórios na armazenagem e promoção de dados em escala global. Essa plataforma permite que pesquisadores compartilhem e preservem seus dados de pesquisa, facilitando o acesso e a reutilização por outros cientistas. Além de ampliar a visibilidade dos dados, o Zenodo assegura sua preservação a longo prazo, garantindo que informações relevantes sejam protegidas contra a perda ao longo do tempo.

A interoperabilidade de repositórios, baseada em padrões como Dublin Core e METS, promove o uso cruzado de dados em escala global. Isso não apenas facilita a colaboração entre pesquisadores de diferentes disciplinas, mas também contribui para avanços interdisciplinares, permitindo que dados de diferentes áreas do conhecimento sejam integrados e utilizados de maneira mais eficaz (Santos; Flores, 2020). Essa abordagem colaborativa é essencial em um mundo onde os desafios científicos são cada vez mais complexos e exigem soluções que vão além das fronteiras disciplinares tradicionais. Dessa forma, os dados abertos, orientados pelos princípios FAIR, tornam-se essenciais para promover a transparência, a colaboração e a inovação na pesquisa científica. A adoção de práticas que garantam o acesso amplo e a reutilização desses dados traz benefícios não apenas para a comunidade científica, mas também para a sociedade em geral, permitindo que informações relevantes sejam empregadas no enfrentamento de desafios globais. Diante das reflexões apresentadas sobre Dados Abertos, algumas direções promissoras para pesquisas futuras podem ser destacadas.

4 <https://zenodo.org/>

5 <https://figshare.com/>

SUMÁRIO

PESQUISA REPRODUTÍVEL ABERTA

A pesquisa reprodutível aberta é um conceito que visa garantir que os estudos científicos sejam verificáveis e confiáveis, promovendo uma cultura de transparência e rigor na pesquisa. A reprodutibilidade é um dos pilares fundamentais da ciência, pois permite que outros pesquisadores verifiquem e validem os resultados obtidos, contribuindo para a credibilidade da pesquisa científica Shintaku *et al.* (2019). A utilização de ferramentas como cadernos de laboratório abertos, exemplificados pelo Jupyter Notebook⁶, é uma prática que tem ganhado destaque nesse contexto. Esses cadernos permitem que os pesquisadores documentem suas metodologias e fluxos de trabalho de maneira detalhada, facilitando a replicação dos resultados por outros cientistas (Silveira *et al.*, 2023).

Na área da neurociência, a prática de compartilhar protocolos experimentais de forma aberta tem sido particularmente benéfica. A disponibilização desses protocolos não apenas facilita a validação de estudos, mas também acelera o avanço do conhecimento ao permitir que diferentes grupos de pesquisa colaborem e construam sobre os resultados uns dos outros (Martins, 2024). Essa abordagem colaborativa é essencial em campos onde a complexidade dos experimentos pode dificultar a replicação, e a transparência nos métodos utilizados é fundamental para a construção de uma base sólida de conhecimento (Costa, 2023).

As metodologias de pesquisa reprodutível aberta incluem investigações qualitativas que exploram as percepções de pesquisadores em relação à reprodutibilidade, bem como análises quantitativas que examinam as taxas de replicação em diferentes áreas do saber. Essas investigações podem revelar não apenas a aceitação e os desafios enfrentados pelos pesquisadores ao adotar práticas de

SUMÁRIO

reprodutibilidade, mas também fornecer dados sobre a eficácia dessas práticas em diferentes contextos (Bertin *et al.*, 2019). Por exemplo, entrevistas com editores de periódicos podem oferecer indicações sobre como a reprodutibilidade é percebida e valorizada na avaliação de manuscritos, enquanto análises de redes de citação podem ajudar a entender como a reprodutibilidade impacta o engajamento acadêmico e a disseminação do conhecimento (Mendes-Da-Silva, 2019).

Além disso, a adoção de práticas de pesquisa reprodutível aberta está alinhada com os princípios da Ciência Aberta, que busca não apenas aumentar a transparência, mas também promover a inclusão e a colaboração entre pesquisadores de diferentes disciplinas e regiões (Santos, 2021). Essa abordagem é especialmente relevante em um momento em que a ciência enfrenta desafios globais, como pandemias e crises ambientais, que exigem uma resposta rápida e colaborativa da comunidade científica. Portanto, a pesquisa reprodutível aberta transcende a preocupação com o rigor científico, constituindo-se também em uma estratégia essencial para promover um ambiente de pesquisa mais colaborativo e atento às demandas sociais. A partir das reflexões realizadas sobre a Pesquisa Reprodutível Aberta, é possível identificar caminhos relevantes para estudos futuros. Dentre as possibilidades destacam-se investigações sobre as barreiras práticas e culturais enfrentadas por diferentes áreas científicas na adoção dessas práticas, bem como análises do impacto da pesquisa reprodutível aberta na confiabilidade e na aceitação dos resultados por parte de comunidades acadêmicas específicas.

AVALIAÇÃO DA CIÊNCIA ABERTA

A avaliação de práticas voltadas à abertura do conhecimento representa uma transformação significativa nos critérios de validação científica, promovendo abordagens mais transparentes

SUMÁRIO

e inclusivas. Esse movimento busca não apenas democratizar o acesso à informação, mas também redefinir os modos pelos quais a pesquisa é reconhecida e valorizada. Um dos principais aspectos desse processo é a adoção de revisões por pares abertas, como as implementadas pela revista *Nature Communications*⁷. Tais revisões promovem maior responsabilidade e colaboração entre os pesquisadores, permitindo que o processo avaliativo se torne mais claro e acessível (Grácio *et al.*, 2020).

As métricas responsáveis, que priorizam o impacto social e a relevância colaborativa, desafiam os indicadores puramente quantitativos que tradicionalmente dominam a avaliação científica. Em vez de se concentrar apenas em fatores como o número de citações ou o fator de impacto de uma revista, as métricas responsáveis consideram como a pesquisa afeta a sociedade e como ela é utilizada por diferentes comunidades (Souza; Aganette, 2020). Essa mudança de foco é fundamental, pois reconhece que o valor da pesquisa não se limita à sua publicação em periódicos de alto impacto, mas também inclui seu potencial para gerar mudanças sociais e contribuir para o bem-estar coletivo (Cabsela, 2023).

Pesquisas sobre a avaliação da Ciência Aberta podem investigar o impacto dessas práticas por meio de entrevistas com editores, análises de redes de citação e estudos sobre revisão por pares aberta, conectando-as ao engajamento acadêmico e ao avanço científico. Essa abordagem permite uma compreensão mais aprofundada de como práticas de abertura estão sendo implementadas e percebidas em diferentes contextos, tanto em escala internacional quanto nacional (Ross-Hellauer, 2017; Silva; Inácio Júnior, 2024). Além disso, a Ciência Aberta promove uma nova forma de colaboração entre pesquisadores, instituições e a sociedade, desafiando as estruturas tradicionais de poder e controle sobre o conhecimento (Silva; Flores, 2018).

SUMÁRIO

A implementação da Ciência Aberta também enfrenta desafios, incluindo a resistência a mudanças nas práticas de avaliação e a necessidade de capacitação dos pesquisadores para adotar essas novas abordagens (Santos; Flores, 2015). No entanto, a crescente aceitação da Ciência Aberta como um modelo viável para a pesquisa científica sugere que, à medida que mais instituições e pesquisadores adotam essas práticas, a avaliação da Ciência Aberta pode se tornar uma norma, promovendo um ambiente de pesquisa mais inclusivo e colaborativo.

Em suma, avaliar práticas de abertura na ciência implica redefinir critérios de validação científica e fomentar uma cultura de transparência e colaboração com potencial de impacto tanto na pesquisa quanto na sociedade. À medida que mais países e instituições adotam esse modelo, espera-se que suas diretrizes sejam progressivamente incorporadas aos processos de avaliação, impulsionando um avanço mais equitativo, acessível e alinhado aos princípios de ciência responsável. A partir das reflexões apresentadas, surgem perspectivas relevantes para futuras investigações. Recomenda-se, por exemplo, explorar o desenvolvimento de novas métricas qualitativas e quantitativas voltadas à mensuração do impacto social dessas práticas, assim como conduzir estudos comparativos sobre a percepção de autores, revisores e leitores frente às diferentes abordagens de revisão por pares abertas.

POLÍTICAS DE CIÊNCIA ABERTA

As políticas de Ciência Aberta são um conjunto de diretrizes institucionais e regulamentos estabelecidos por financiadores e organizações que visam incentivar práticas de abertura na pesquisa científica. Essas políticas têm como objetivo promover a transparência, a colaboração e a acessibilidade dos dados e resultados de

SUMÁRIO

pesquisa, contribuindo para um ambiente científico mais inclusivo e inovador. Um exemplo significativo dessas políticas é o programa *Horizon Europe*⁸, da União Europeia, que exige que os dados gerados a partir de projetos financiados sejam disponibilizados em repositórios abertos.

O Horizon Europe não apenas estabelece a obrigatoriedade da disponibilização de dados, mas também orienta os pesquisadores sobre como gerenciar e compartilhar esses dados de maneira eficaz. Essa abordagem é fundamental para garantir que os resultados da pesquisa possam ser utilizados e reutilizados por outros pesquisadores, promovendo a colaboração interdisciplinar e a inovação. Além disso, a exigência de que os dados sejam depositados em repositórios abertos ajuda a garantir a rastreabilidade e a integridade dos dados, aspectos essenciais para a credibilidade da pesquisa científica.

As políticas de Ciência Aberta também desempenham um papel essencial na harmonização de práticas globais, fortalecendo a adesão à Ciência Aberta em diferentes contextos. Ao estabelecer diretrizes claras e padronizadas, essas políticas facilitam a comparação e a colaboração entre pesquisadores de diferentes países e instituições, promovendo um ecossistema de pesquisa mais coeso e colaborativo. Essa padronização é particularmente importante em um mundo cada vez mais globalizado, onde a pesquisa muitas vezes transcende fronteiras nacionais e disciplinares.

Para avaliar a eficácia dessas políticas, abordagens qualitativas podem ser utilizadas. Essas abordagens podem incluir entrevistas com pesquisadores e gestores, análise de documentos e observação de práticas institucionais. A pesquisa qualitativa permite uma compreensão mais profunda de como as políticas de Ciência Aberta estão sendo implementadas na prática e quais desafios e oportunidades surgem nesse processo. Além disso, estudos comparativos

SUMÁRIO

podem ser realizados para identificar lacunas nas políticas existentes e oportunidades para melhorias institucionais. Esses estudos podem comparar diferentes políticas de Ciência Aberta em diferentes contextos, analisando suas características, impactos e resultados.

De modo geral, as políticas voltadas à abertura da ciência são fundamentais para promover a transparência, a colaboração e a acessibilidade na pesquisa científica. Por meio de diretrizes institucionais e orientações de agências de fomento, como o *Horizon Europe*, essas iniciativas incentivam práticas que beneficiam não apenas os pesquisadores, mas também a sociedade como um todo. A análise qualitativa e comparativa desses marcos pode contribuir significativamente para aprimorar sua implementação e eficácia, favorecendo um futuro mais aberto, colaborativo e inclusivo para a ciência. Com base nas reflexões realizadas sobre esse conjunto de políticas, algumas oportunidades relevantes para investigações futuras podem ser destacadas. Sugere-se, por exemplo, que estudos posteriores realizem análises comparativas da eficácia dessas estratégias em diferentes países e contextos institucionais, identificando boas práticas, desafios frequentes e possíveis soluções para sua consolidação.

FERRAMENTAS DE CIÊNCIA ABERTA

As ferramentas digitais são indispensáveis para implementar práticas de Ciência Aberta. Plataformas colaborativas, como o Open Science Framework, permitem que os pesquisadores compartilhem dados e resultados em tempo real, promovendo um ambiente de trabalho mais integrado e eficiente (Lui *et al.* (2022); Gesing *et al.*, (2022). Essas plataformas não apenas conectam cientistas de diferentes áreas, mas também incentivam a troca de ideias e a coautoria, o que é fundamental para enfrentar desafios complexos, como as mudanças climáticas e as pandemias (Gesing *et al.*, 2022).

SUMÁRIO

Identificadores persistentes, como os Digital Object Identifiers (DOIs), são essenciais para assegurar a rastreabilidade de publicações e dados. Eles garantem que os pesquisadores possam localizar e citar corretamente os trabalhos de outros, o que é vital para a integridade da pesquisa científica (Barbosa, 2023). Além disso, padrões de metadados, como o Dublin Core, facilitam a interoperabilidade entre diferentes repositórios e plataformas, permitindo que os dados sejam facilmente compartilhados e reutilizados (Grácio, 2018). Essa interoperabilidade é crucial para a construção de um ecossistema de pesquisa mais coeso e acessível.

Estudos experimentais têm mostrado que a utilização dessas ferramentas digitais pode impactar positivamente a produtividade dos pesquisadores, o compartilhamento de dados e a colaboração científica. Por exemplo, a análise de redes de colaboração pode revelar como as interações entre pesquisadores de diferentes instituições e disciplinas podem levar a inovações significativas e a um aumento na qualidade da pesquisa (Ferreira, 2023). Além disso, a adoção de ferramentas de Ciência Aberta pode melhorar a visibilidade e o impacto das publicações científicas, contribuindo para um maior reconhecimento do trabalho realizado (Nascimento, 2023).

A Ciência Aberta vai além da simples disponibilização de dados e publicações, envolvendo práticas transparentes ao longo de todo o processo científico, da coleta de informações à divulgação dos resultados (Faria; Gadelha, 2016). Isso inclui o uso de ferramentas digitais que favorecem a colaboração e a reprodutibilidade, como softwares de análise e plataformas de compartilhamento (Lacerda; Silva, 2015).

Nesse contexto, tais recursos são essenciais para tornar a pesquisa mais inclusiva, acessível e conectada aos desafios contemporâneos. A continuidade no desenvolvimento e na adoção dessas tecnologias é estratégica para fortalecer práticas abertas e colaborativas. A partir das reflexões sobre esse tema, emergem

SUMÁRIO

questões ainda pouco exploradas, apontando caminhos para futuras investigações. Sugere-se, por exemplo, a realização de estudos que avaliem o impacto dessas ferramentas na produtividade científica e na colaboração interdisciplinar, com base em análises empíricas que revelem como diferentes comunidades científicas vêm incorporando essas soluções.

EDUCAÇÃO ABERTA

A Educação Aberta é um conceito que visa promover a capacitação e a democratização do conhecimento por meio de Recursos Educacionais Abertos (REA) e currículos baseados em Ciência Aberta. A UNESCO tem desempenhado um papel fundamental nesse movimento, liderando iniciativas que buscam criar materiais educacionais acessíveis globalmente, ampliando assim o alcance da educação científica (Seco; Mendes, 2016); Zanin, 2017). Os REA são definidos como materiais educacionais que podem ser utilizados, adaptados e redistribuídos livremente, o que estimula uma abordagem colaborativa e inclusiva na educação (Martins *et al.*, 2022; Matias *et al.*, 2016).

A pesquisa-ação é uma metodologia que pode ser utilizada para desenvolver e avaliar programas educacionais, permitindo uma análise aprofundada de como esses programas transformam práticas pedagógicas e fomentam a adoção da Ciência Aberta (Netto *et al.*, 2023; Oliveira *et al.*, 2015). Essa abordagem envolve a participação ativa de educadores e alunos no processo de pesquisa, promovendo um ciclo contínuo de reflexão e melhoria das práticas educacionais (Campos; Hashizume, 2022; Silveira *et al.*, 2023). A pesquisa-ação, além de contribuir para a formação de professores, favorece também a criação de ambientes de aprendizagem mais inclusivos e adaptados às necessidades dos alunos (Brostolin; Souza, 2023).

SUMÁRIO

Além disso, a inclusão de aspectos de acessibilidade nos REA é um desafio que ainda precisa ser enfrentado. Embora haja um crescente número de iniciativas que disponibilizam recursos educacionais abertos, a acessibilidade continua a ser uma preocupação central para garantir que todos os aprendizes possam se beneficiar desses materiais (Martins *et al.*, 2022; Camillo *et al.*, 2019). A promoção de uma educação verdadeiramente aberta requer um compromisso com a inclusão, assegurando que os recursos sejam utilizáveis por todos, independentemente de suas habilidades ou contextos.

A interseção entre educação aberta e pesquisa-ação também pode ser vista como uma oportunidade para inovar nas práticas pedagógicas. Ao integrar a pesquisa-ação com a criação e o uso de REA, educadores podem explorar novas formas de engajamento e colaboração, resultando em experiências de aprendizagem mais significativas e impactantes (Franzen; Lunardi, 2019; Fontoura; Procasko, 2019). Essa abordagem não apenas enriquece o processo de ensino-aprendizagem, mas também fortalece a comunidade educacional, promovendo um ambiente de aprendizado contínuo e adaptativo.

Desse modo, a Educação Aberta, apoiada por iniciativas da UNESCO e pela metodologia da pesquisa-ação, representa uma estratégia para democratizar o conhecimento e promover a inclusão educacional. O desenvolvimento de Recursos Educacionais Abertos (REA) acessíveis e a adoção de práticas pedagógicas inovadoras são fundamentais para assegurar que a educação responda adequadamente às demandas de todos os aprendizes, contribuindo diretamente para a construção de uma sociedade mais justa, inclusiva e equitativa. Considerando as perspectivas apresentadas sobre Educação Aberta, algumas áreas emergem como especialmente relevantes para futuras investigações acadêmicas. Recomenda-se, por exemplo, explorar a criação e avaliação de REA acessíveis voltados a públicos diversos, realizando estudos que examinem tanto a eficácia pedagógica desses materiais quanto sua adoção em diferentes contextos educacionais.

LICENCIAMENTO ABERTO

O conceito de Licenciamento Aberto, particularmente através de licenças como as da Creative Commons (CC), desempenha um papel muito importante na democratização do acesso ao conhecimento científico. Essas licenças estabelecem os termos de uso e reutilização de conteúdos, permitindo que pesquisadores e o público em geral acessem e compartilhem informações de maneira legal e ética. A licença CC-BY, por exemplo, permite a redistribuição de artigos científicos desde que a devida atribuição ao autor original seja feita, promovendo assim a circulação do conhecimento sem as barreiras tradicionais impostas pelos direitos autorais.

A adoção de licenças abertas tem sido incentivada por diversas instituições e políticas de acesso aberto no Brasil, que visam regulamentar e fortalecer a prática de publicação em periódicos e o depósito em repositórios digitais (Café *et al.*, 2022; Furnival *et al.*, 2015). Essas políticas não apenas refletem um compromisso com a transparência e a acessibilidade, mas também servem como instrumentos de governança epistêmica, moldando a forma como o conhecimento é produzido e disseminado (Café *et al.*, 2022). A análise das políticas institucionais revela que, apesar do crescimento do acesso aberto, ainda existem desafios significativos, como a resistência de algumas editoras e a necessidade de financiamento para cobrir as taxas de processamento de artigos (Article Processing Charges - APC) (Pavan; Barbosa, 2017; Appel; Albagli, 2019).

Metodologicamente, a pesquisa sobre licenciamento aberto pode incluir entrevistas com gestores de periódicos e análises quantitativas que avaliam o impacto de diferentes modelos de licenciamento na disseminação do conhecimento. Estudos indicam que a publicação em acesso aberto não só aumenta a visibilidade dos artigos, mas também pode resultar em um maior número de citações, um fenômeno conhecido como “Open Access Citation Advantage”

SUMÁRIO

(Cintra *et al.*, 2017). Além disso, a análise de dados sobre a produção científica em periódicos de acesso aberto mostra um crescimento significativo, refletindo uma mudança nas preferências dos pesquisadores em relação ao compartilhamento de suas obras (Caballero-Rivero *et al.*, 2019).

A importância do licenciamento aberto é ainda mais evidente quando se considera o papel que ele desempenha na promoção da Ciência Aberta. A Ciência Aberta busca não apenas a disponibilização de artigos, mas também a inclusão de dados e materiais de pesquisa, permitindo uma colaboração mais ampla e a reutilização de informações (Fontenelle; Sarti, 2020). Essa abordagem é particularmente relevante em um contexto onde a velocidade da informação e a necessidade de colaboração interdisciplinar são cada vez mais exigidas para enfrentar desafios globais complexos.

Portanto, o licenciamento aberto, através de iniciativas como as licenças Creative Commons, representa uma estratégia fundamental para a promoção do acesso ao conhecimento científico, equilibrando a necessidade de proteção dos direitos autorais com a urgência de democratizar a informação. A continuidade no desenvolvimento de políticas que incentivem essa prática é essencial para assegurar que o conhecimento científico esteja amplamente disponível e beneficie toda a sociedade. Considerando os aspectos abordados sobre Licenciamento Aberto, é possível destacar algumas oportunidades significativas para futuras investigações acadêmicas. Sugere-se, por exemplo, que estudos posteriores analisem comparativamente os impactos gerados por diferentes modelos de licenciamento aberto sobre o uso e a reutilização do conhecimento científico, identificando práticas bem-sucedidas e desafios legais encontrados em diversos contextos institucionais.

PRESERVAÇÃO DIGITAL

A preservação digital é um componente essencial no contexto das práticas abertas, pois garante que os dados e os resultados de pesquisa gerados sejam acessíveis e utilizáveis a longo prazo. A Ciência Aberta, que promove a transparência, a colaboração e o acesso livre ao conhecimento científico, depende da preservação digital para assegurar que os dados gerados em projetos de pesquisa, incluindo aqueles de ciência cidadã, permaneçam disponíveis para futuras investigações e aplicações práticas (Silva, 2023; Ribeiro et al., 2022).

A preservação digital envolve a implementação de estratégias e práticas que asseguram a integridade, acessibilidade e usabilidade dos dados ao longo do tempo. Isso é particularmente relevante em um contexto onde a quantidade de dados gerados está aumentando exponencialmente, e onde a digitalização de informações científicas se torna a norma (Manghi *et al.*, 2020; Meghini, 2013). A adoção de padrões de metadados, como o PREMIS (Preservation Metadata: Implementation Strategies), é essencial para a gestão eficaz dos dados digitais e para a sua preservação a longo prazo (Castro; Alves, 2021). Esses padrões ajudam a garantir que os dados sejam documentados de forma adequada, facilitando sua recuperação e reutilização por outros pesquisadores.

Além disso, a infraestrutura de preservação digital deve ser robusta e alinhada com normas reconhecidas, como a ISO 16363, que estabelece critérios para repositórios digitais confiáveis (Santos; Flores, 2020; Benin *et al.*, 2022). A segurança e a gestão de riscos são aspectos críticos que precisam ser considerados na construção de repositórios digitais, garantindo que os dados estejam protegidos contra perda ou corrupção (Santos; Flores, 2020; Patterson, 2021). A implementação de serviços em nuvem também tem se mostrado uma solução viável para a preservação digital, permitindo que instituições com recursos limitados possam armazenar e gerenciar seus dados de forma eficaz (Matusiak *et al.*, 2017).

SUMÁRIO

A preservação digital não se limita apenas à manutenção de dados; ela também envolve a curadoria digital, que se refere ao gerenciamento ativo de dados ao longo de seu ciclo de vida, desde a criação até a preservação (Pasqui, 2024). Isso inclui a aplicação de práticas de “preservação digital por design”, onde as considerações de preservação são integradas desde o início de qualquer projeto de pesquisa (Pasqui, 2024). Essa abordagem proativa é essencial para assegurar que os dados sejam preservados de maneira eficaz, garantindo seu acesso e reutilização pelas futuras gerações de pesquisadores.

Considerando os desafios anteriormente apresentados sobre Preservação Digital, surgem importantes oportunidades para aprofundar esse tema em pesquisas futuras. Recomenda-se, por exemplo, investigar a eficácia de diferentes estratégias de preservação digital, especialmente em contextos institucionais com recursos financeiros ou tecnológicos limitados.

Além disso, estudos adicionais poderiam examinar o impacto de tecnologias emergentes, como os serviços baseados em nuvem, sobre a preservação e gestão de dados científicos, identificando benefícios, limitações e desafios para sua implementação prática.

INOVAÇÃO ABERTA

A inovação aberta é um conceito que se destaca por sua capacidade de conectar ciência, sociedade e economia, promovendo colaborações entre os setores público e privado. Essa abordagem não apenas estimula a criatividade e a troca de conhecimentos, mas também facilita a criação de soluções inovadoras para desafios globais, como demonstrado pela Open COVID Pledge⁹. Esta iniciativa

SUMÁRIO

incentivou empresas a liberar patentes para combater a pandemia, ilustrando como a abertura pode gerar respostas rápidas e eficazes em momentos de crise (Carvalho & Sugano, 2016; Brüggemann *et al.*, 2022). A colaboração entre diferentes atores, incluindo universidades, empresas e governos, é fundamental para o sucesso da inovação aberta, pois permite a combinação de recursos e expertise diversificados (Vasconcelos, 2023; Souza *et al.*, 2022).

Além disso, metodologias como a análise de redes são essenciais para mapear parcerias e avaliar seu impacto na criação de novas tecnologias e práticas. A análise de redes permite identificar e entender as interações entre os diversos participantes do ecossistema de inovação, facilitando a identificação de oportunidades de colaboração e a maximização dos resultados (Mazzetti *et al.*, 2019; Cândido, 2017). A inovação aberta não se limita a um único setor; ela pode ser adaptada a diferentes contextos, como demonstrado em estudos sobre companhias de saneamento, onde a gestão da propriedade intelectual e a pesquisa e desenvolvimento desempenham papéis centrais (Vasconcelos, 2023).

A importância da inovação aberta é ainda mais evidente quando se considera que ela pode levar a uma maior agilidade organizacional e a um aumento na criação de produtos inovadores. As empresas que adotam essa abordagem frequentemente experimentam custos reduzidos e a capacidade de explorar novos mercados, contrastando com a inovação fechada, que tende a focar em melhorias incrementais (Brüggemann *et al.*, 2022; Cândido, 2017). Assim, a inovação aberta não apenas promove a eficiência, mas também a eficácia na resposta a desafios complexos, como os enfrentados durante a pandemia de COVID-19, onde a colaboração entre os setores público e privado foi vital para a expansão da capacidade de atendimento e a equidade no acesso aos serviços de saúde (Costa *et al.*, 2020; Santos *et al.*, 2022).

SUMÁRIO

Portanto, a inovação aberta representa uma estratégia possibilitando que diferentes setores trabalhem juntos em prol de soluções inovadoras e sustentáveis. Diante do contexto apresentado sobre Inovação Aberta, surgem questões relevantes que merecem atenção em estudos futuros, tanto acadêmicos quanto profissionais.

Uma direção importante para pesquisas adicionais consiste em investigar formas efetivas de implementar a inovação aberta em diferentes setores econômicos e sociais. Nesse sentido, é particularmente valioso explorar como parcerias estratégicas entre universidades, setor privado e sociedade civil podem potencializar resultados positivos e fomentar práticas inovadoras sustentáveis.

ESTRUTURANDO A INVESTIGAÇÃO EM CIÊNCIA ABERTA: QUESTÕES, ABORDAGENS E MÉTODOS

Após a exploração das diversas facetas da Ciência Aberta, esta seção final atua como uma síntese propositiva e um guia metodológico. O objetivo é organizar as oportunidades de investigação em três pontos centrais, apresentando não apenas as questões-chave, mas também detalhando as abordagens e os métodos mais adequados para respondê-las, consolidando um roteiro para o desenvolvimento de novas pesquisas no campo.

SUMÁRIO

PONTO 1: PRÁTICAS E METODOLOGIAS ABERTAS

Este ponto foca no “como fazer”, investigando a implementação, o uso e o impacto de ferramentas e práticas colaborativas que sustentam o dia a dia da Ciência Aberta.

QUESTÕES-CHAVE:

Como as novas ferramentas e plataformas digitais (ex: cadernos abertos, repositórios de dados, frameworks colaborativos) estão efetivamente transformando as rotinas de pesquisa, a cultura de colaboração e a reprodutibilidade dos resultados em diferentes campos do saber?

De que maneira projetos de ciência cidadã podem ser desenhados para maximizar a inclusão social e a equidade, ao mesmo tempo que garantem a qualidade e a confiabilidade dos dados gerados por participantes não especialistas?

Quais abordagens pedagógicas e Recursos Educacionais Abertos (REA) são mais eficazes para a formação de competências em Ciência Aberta em diferentes níveis de ensino, desde a graduação até a formação continuada de pesquisadores, fomentando uma nova cultura científica?

Como podem ser estruturadas colaborações eficazes entre universidades, empresas e a sociedade civil, sob os princípios da Inovação Aberta, para acelerar a tradução do conhecimento científico em soluções práticas e de valor social?

SUMÁRIO

ABORDAGENS METODOLÓGICAS RECOMENDADAS:

Para investigar essas questões, as seguintes abordagens são particularmente eficazes:

- **Estudos de Caso:** Constituem uma abordagem valiosa para investigar práticas de Ciência Aberta em contextos reais, possibilitando uma compreensão profunda de como os princípios da abertura são implementados e seus impactos práticos. A metodologia envolve selecionar casos representativos, como universidades com políticas robustas de dados abertos, revistas científicas que adotam práticas inovadoras de revisão aberta, ou projetos de ciência cidadã, garantindo relevância e profundidade. A análise pode incluir documentos institucionais, relatórios, observação direta e entrevistas, permitindo identificar mecanismos de incentivo, desafios culturais e soluções práticas.
- **Pesquisa-Ação Participativa (PAP):** É especialmente relevante para iniciativas que conectam ciência e sociedade, como projetos de ciência cidadã ou educação científica comunitária. Nesse método, pesquisadores e membros das comunidades colaboram desde a definição dos problemas até a aplicação prática dos resultados. Um projeto pode envolver comunidades locais no monitoramento ambiental, utilizando oficinas e rodas de conversa para analisar os dados coletivamente. A PAP fortalece o vínculo entre conhecimento científico e saber comunitário, garantindo práticas científicas mais inclusivas e relevantes socialmente.
- **Análise Qualitativa de Percepções e Atitudes:** Esta abordagem investiga como pesquisadores, gestores e técnicos compreendem e respondem à Ciência Aberta. Utilizando entrevistas semiestruturadas ou grupos focais, com apoio de softwares como o NVivo para análise, é possível captar

SUMÁRIO

preocupações, resistências e aceitação frente a novas ferramentas e práticas. Uma pesquisa pode revelar receios sobre compartilhamento de dados, como plágio ou perda de vantagem competitiva, ajudando a identificar estratégias institucionais eficazes para superar barreiras culturais e aumentar o engajamento.

PONTO 2: MÉTRICAS, AVALIAÇÃO E IMPACTO

Este ponto aborda o desafio de medir e valorizar a produção científica em um ecossistema aberto, questionando os modelos tradicionais de avaliação e buscando novas formas de aferir relevância.

QUESTÕES-CHAVE:

- Quais novos indicadores e métricas responsáveis (qualitativas e quantitativas) podem ser desenvolvidos para avaliar de forma mais justa o impacto acadêmico, social e econômico da pesquisa realizada sob os princípios da Ciência Aberta, indo além das citações em periódicos?
- Qual é o real impacto dos diferentes modelos de Acesso Aberto (Dourado, Verde, Diamante) na visibilidade, no uso e na reutilização da produção científica, especialmente quando se consideram assimetrias entre diferentes áreas do conhecimento e regiões geográficas?

SUMÁRIO

ABORDAGENS METODOLÓGICAS RECOMENDADAS:

Para mensurar o impacto e a eficácia dessas práticas, destacam-se os seguintes métodos:

- **Análise Bibliométrica e de Redes:** É uma abordagem quantitativa essencial para mapear a influência da Ciência Aberta sobre padrões de publicação e colaboração. Utilizando bases como Scopus, Web of Science e Dimensions (muitas vezes com extração de dados via APIs), os pesquisadores coletam dados sobre publicações, citações e redes de coautoria. Softwares especializados como VOSviewer e Gephi permitem criar visualizações interativas que ilustram padrões onde práticas abertas impulsionam colaborações e disseminação de conhecimento.
- **Análise de Impacto:** Busca mensurar os efeitos das práticas abertas em diferentes dimensões (acadêmica, social, econômica). Essa metodologia combina métricas quantitativas, como downloads e citações (altmetrics), com dados qualitativos obtidos por entrevistas ou relatos de usuários. Um estudo pode analisar o impacto de um repositório institucional antes e depois de sua implementação, medindo mudanças em indicadores e identificando efeitos diretos (visibilidade) e indiretos (influência em políticas públicas).
- **Análise Econômica dos Custos e Benefícios:** Avalia os investimentos necessários para implementar práticas de Ciência Aberta e os retornos gerados. A metodologia inclui o levantamento de custos com infraestrutura, treinamento e manutenção, e a quantificação dos benefícios, como a economia gerada por reutilizar dados abertos em comparação à coleta original. Essa abordagem proporciona clareza sobre o valor econômico da abertura científica, orientando decisões estratégicas.

SUMÁRIO

- **Métodos Experimentais:** Testam intervenções específicas em contextos controlados para avaliar causalidade. Por exemplo, pode-se criar um experimento com grupos de controle para testar se artigos com licenças abertas e dados disponíveis têm maior reutilização e credibilidade. Com variáveis controladas e análises estatísticas rigorosas (utilizando softwares como R ou SPSS), essa metodologia oferece evidências robustas para a tomada de decisões políticas e institucionais.

PONTO 3: GOVERNANÇA, ÉTICA E POLÍTICAS

Este ponto concentra-se nas estruturas que regulam, financiam e sustentam a Ciência Aberta, investigando as dimensões políticas, legais, éticas e econômicas do movimento.

QUESTÕES-CHAVE:

Como as políticas institucionais e de agências de fomento estão sendo implementadas e qual sua eficácia para incentivar (ou inibir) a adesão dos pesquisadores às práticas de abertura de dados e publicações?

De que forma os Planos de Gestão de Dados (PGDs) podem ser mais do que instrumentos burocráticos, tornando-se ferramentas estratégicas para aprimorar a qualidade, a interoperabilidade (princípios FAIR) e a preservação digital de longo prazo dos dados de pesquisa?

SUMÁRIO

Quais são os principais dilemas éticos e legais que emergem do compartilhamento de dados em larga escala, incluindo a gestão da propriedade intelectual por meio de licenças abertas, a proteção da privacidade e a soberania dos dados?

Que modelos de sustentabilidade econômica são mais viáveis para garantir a infraestrutura de longo prazo da Ciência Aberta (repositórios, periódicos diamante, etc.) sem criar novas formas de exclusão ou aprofundar desigualdades existentes?

ABORDAGENS METODOLÓGICAS RECOMENDADAS:

A análise da governança e de seus desafios pode ser realizada por meio de:

- **Análise de Políticas Comparadas:** Examina diferentes abordagens institucionais e governamentais para a implementação da Ciência Aberta. A metodologia envolve selecionar documentos normativos relevantes de diferentes países ou instituições para comparar estratégias de financiamento, incentivos e apoio técnico. Essa análise permite identificar boas práticas e avaliar como podem ser adaptadas a contextos variados.
- **Análise Documental:** Envolve o exame sistemático de políticas, regulamentos, relatórios institucionais e termos de licenciamento que orientam a Ciência Aberta. Documentos são selecionados com base em critérios claros e analisados para identificar padrões, inconsistências e avanços. O uso de softwares de análise qualitativa pode facilitar a interpretação dos discursos e práticas consolidadas, contribuindo para a formulação de políticas mais coerentes.

SUMÁRIO

DO PROBLEMA À DECISÃO: UM QUADRO-SÍNTESE PARA A PESQUISA EM CIÊNCIA ABERTA

Para consolidar e operacionalizar a discussão apresentada, esta seção final oferece um quadro-síntese. O objetivo é ilustrar, por meio de exemplos práticos, como questões de pesquisa específicas sobre as facetas da Ciência Aberta podem ser abordadas metodologicamente. A tabela a seguir conecta uma pergunta-tipo a um método de investigação, a exemplos de métricas que podem ser coletadas e a uma potencial decisão de governança que poderia ser informada pelos resultados, encapsulando o ciclo completo da pesquisa orientada para a ação.

Faceta	Pergunta-Tipo	Método Sugerido	Métricas de Exemplo	Decisão de Governança Associada
Acesso Aberto	Acordos transformativos elevam a taxa de abertura sem inflar os custos de publicação (APCs)?	Séries temporais + Análise de Diferença em Diferenças (Diff-in-Diff)	% de artigos em OA por via; APC mediana; tempo até o depósito; distribuição de licenças.	Renovar, ajustar ou encerrar o acordo; criar um fundo para APCs ou para o Acesso Diamante.
Dados Abertos (FAIR)	Quais fatores em um dataset (metadados, formato, etc.) predizem sua reutilização por outros pesquisadores?	Regressão multivariada + Auditoria de maturidade FAIR	Nível de maturidade FAIR; atribuição de DOI/DataCite; número de citações do dataset; downloads.	Exigir Plano de Gestão de Dados (PGD); definir padrão mínimo de metadados; política de anonimização.
Reprodutibilidade	A exigência de checklists e o uso de containers (ex: Docker) aumentam as taxas de reprodução de estudos?	Ensaio por Clusters Randomizado em periódicos selecionados	% de artigos com código/ ambiente disponível; taxa de sucesso na reprodução por pares.	Implementar mandato de código aberto; adotar padrão de "empacotamento" (container) para submissões.
Avaliação (Peer Review)	A revisão por pares aberta altera o tempo do processo, a qualidade dos pareceres ou a decisão final do editor?	Quase-experimento (comparando periódicos com e sem revisão aberta)	Tempo do ciclo de revisão; taxa de aceitação/rejeição; número e perfil dos revisores.	Adotar política de revisão por pares aberta; desenvolver métricas responsáveis para avaliação.

SUMÁRIO

Políticas de C.A.	Quais cláusulas em políticas de fomento (ex: obrigatoriedade, exceções) elevam a conformidade dos pesquisadores?	Estudo comparativo multinível + Análise textual de documentos	Taxa de depósito em repositórios; % de PGDs submetidos; tipos de exceções mais justificadas.	Reformular cláusulas da política; definir sanções, incentivos ou mecanismos de apoio.
Ciência Cidadã	A participação em projetos de ciência cidadã eleva o letramento científico e o engajamento cívico dos voluntários?	Pesquisa-Ação Participativa (PAP) + Survey (pré e pós-teste)	Ganho de conhecimento validado; taxa de retenção de voluntários; qualidade dos dados coletados.	Desenvolver protocolos de capacitação para voluntários; estabelecer diretrizes de ética e consentimento.
Preservação Digital	Uma infraestrutura de nuvem híbrida reduz os riscos de perda de dados e os custos a longo prazo em comparação com servidores locais?	Avaliação baseada na ISO 16363 + Análise de custos	Taxa de integridade (checksums); detecção de <i>bit rot</i> ; custo total por TB/ano.	Adotar política de "preservação por design" (<i>by design</i>); definir orçamento e papéis institucionais.

Este quadro não é exaustivo, mas serve como um roteiro e um ponto de partida, demonstrando a relação intrínseca entre escolhas metodológicas rigorosas, a geração de evidências (métricas) e a capacidade de subsidiar decisões estratégicas (governança). Ele materializa o potencial da pesquisa sobre a Ciência Aberta para não apenas compreender, mas também transformar o próprio ecossistema científico, fechando o ciclo que vai do questionamento à ação informada.

REFERÊNCIAS

APPEL, J.; ALBAGLI, S. The adoption of Article Processing Charges as a business model by Brazilian Open Access journals. **Transinformação**, v. 31, e180045, 2019. doi:10.1590/2318-0889201931e180045.

BARBOSA, A. Materiais didáticos digitais para o ensino/aprendizagem das ciências naturais: uma análise bibliométrica. **Texto Livre: Linguagem e Tecnologia**, v. 16, n. 1, p. 1-20, 2023. doi:10.1590/1983-3652.2023.46865.

SUMÁRIO

BENIN, *et al.* Digital open repositories: reliability evaluation based on ISO 16363 criteria. **ANIS**, 2022. doi:10.47909/anis.978-9916-9760-3-6.90.

BERTIN, J. R.; CUNHA, M. A.; GONÇALVES, M. A. A parceria para Governo Aberto como plataforma para o avanço da Ciência Aberta no Brasil. **Transinformação**, v. 11, n. 1, p. 1-12, 2019. doi:10.1590/2318-0889201931e190020.

BROSTOLIN, R.; SOUZA, L. Educação Infantil e Educação Especial: implicações para a formação de professores em espaços inclusivos. **Série-Estudos - Periódico do Programa de Pós-Graduação em Educação da UCDB**, v. 28, n. 62, 2023. doi:10.20435/serieestudos.v28i62.1687.

BRÜGGEMANN, L.; *et al.* Influência do sistema de mensuração de desempenho na agilidade organizacional e na inovação aberta. **Revista de Contabilidade e Organizações**, 2022. doi:10.11606/issn.1982-6486.rco.2022.193897.

CABALLERO-RIVERO, F.; GONÇALVES, R.; CAFE, R. Práticas de Ciência Aberta da comunidade acadêmica brasileira: estudo a partir da produção científica. **Transinformação**, v. 31, e190029, 2019. doi:10.1590/2318-0889201931e190029.

CABSELA, C. P.; MACIE, G. C. Políticas de preservación digital: mapeo de aspectos legales en Brasil y Mozambique. **Revista EDICIC**, v. 2, n. 1, 2023. Disponível em: <https://doi.org/10.62758/re.v2i1.150>. Acesso em: 3 jan. 2025.

CAMILLO, T.; *et al.* Pesquisa-Ação como Resposta Metodológica aos Desafios de Design Social. **PED**, 2019. doi:10.5151/ped2018-3.2_aco_23.

CAMPOS, L.; HASHIZUME, J. Políticas Educacionais. **Cadernos do GPOSHEE On-line**, v. 6, n. 2, 2022. doi:10.33241/cadernosdogposhee.v6i2.9199.

CÂNDIDO, R. Gestão da Informação e Inovação Aberta. **Brazilian Journal of Information Science**, 2017. doi:10.36311/1981-1640.2017v11n2.07p72.

CASTRO, A.; ALVES, R. Cloud services e o padrão PREMIS. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, 2021. doi:10.20396/rdbci.v19i0.8661384.

CAFÉ, R.; ALBORNOZ, L.; GONÇALVES, R. Análise de domínio de políticas institucionais de acesso aberto no Brasil. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 20, 2022. doi:10.20396/rdbci.v20i00.8670092.

SUMÁRIO

CINTRA, R.; RIBEIRO, A.; BARBOSA, M. Vantagens de citação do acesso aberto em periódicos selecionados da Ciência da Informação: uma análise ampliada aos indicadores altmétricos. **Informação & Informação**, v. 22, n. 1, p. 129-144, 2017. doi:10.5433/1981-8920.2017v22n1p129.

COSTA, R. Ciência Aberta na Revista REAMEC: políticas e método de submissão e avaliação de manuscritos. **REAMEC**, v. 11, n. 1, 2023. doi:10.26571/reamec.v11i1.16911.

EISMANN, M. T. Open access. **Optical Engineering**, v. 56, n. 12, p. 1, 2017. DOI: 10.1117/1.oe.56.12.120101.

FARIA, A.; GADELHA, C. A Gestão da Informação na Biblioteca Universitária como ferramenta auxiliar ao planejamento estratégico: um exemplo de análise de impacto. **Revista Conhecimento em Ação**, v. 1, n. 1, p. 1-10, 2016. doi:10.47681/rca.v1i1.3372.

FECHER, Benedikt; FRIESIKE, Sascha. Open science: one term, five schools of thought? In: BARTLING, Sönke; FRIESIKE, Sascha (ed.). **Opening science: the evolving guide on how the web is changing research, collaboration and scholarly publishing**. Cham: Springer, 2014. p. 17-47. DOI: 10.1007/978-3-319-00026-8_2.

FERREIRA, R. Publicação Científica, Ciência Aberta e Avaliação de Investigação — Breve Análise Crítica. **Revista de Educação e Pesquisa em Ciências**, v. 6, n. 1, p. 66-78, 2023. doi:10.21814/uminho.ed.66.6.

FONTENELLE, L.; SARTI, T. Acesso aberto a artigos, dados e materiais de pesquisa na RBMFC. **Revista Brasileira de Medicina de Família e Comunidade**, v. 15, n. 42, p. 2671, 2020. doi:10.5712/rbmfc15(42)2671.

FONTOURA, J.; PROCASKO, P. O Mestrado Profissional em Informática na Educação do IFRS: Temas de Pesquisa, Objetos de Investigação e Produtos Privilegiados. **Renote**, 2019. doi:10.22456/1679-1916.99533.

FRANZEN, E.; LUNARDI, L. O Projeto Político Pedagógico da Educação Infantil: Uma proposta de construção dialógica na Rede Municipal de Ensino de Toropi-RS. **Educação por Escrito**, 2019. doi:10.15448/2179-8435.2019.1.32383.

FURNIVAL, A.; RODRIGUES, A.; RIBEIRO, A. As políticas de direitos autorais e de reuso presentes nas revistas brasileiras de acesso aberto das áreas biológicas e de saúde disponibilizadas na plataforma SciELO-Brasil. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 20, n. 44, p. 25-38, 2015. doi:10.5007/1518-2924.2015v20n44p25.

SUMÁRIO

GESING, S.; MURRAY, J.; *et al.* Open Science via HUBzero: Exploring Five Science Gateways Supporting and Growing their Open Science Communities. In: HICSS. Proceedings of the 55th Hawaii International Conference on System Sciences. 2022. doi:10.24251/hicss.2022.090.

GRÁCIO, Maria Claudia Cabrini. Colaboração científica: indicadores relacionais de coautoria. **Brazilian Journal of Information Science: research trends**, Marília, SP, v. 12, n. 2, 2018. DOI: 10.36311/1981-1640.2018.v12n2.04.p24. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/7976>. Acesso em: 3 jan. 2025.

GRÁCIO, J. C. A.; TROITIÑO, S.; MÁDIO, T. C. d. C.; BREGA, J. R. F.; MORAES, M. B. Modelo para elaboração de políticas de preservação digital de documentos de arquivo por instituições de ensino superior: o caso da Unesp. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 14, n. 3, 2020. Disponível em: <https://doi.org/10.29397/reciis.v14i3.2111>. Acesso em: 3 jan. 2025.

JOHNS HOPKINS UNIVERSITY. Coronavirus Resource Center. Baltimore: Johns Hopkins University, 2020. Disponível em: <https://coronavirus.jhu.edu/>. Acesso em: 29 mar. 2025.

LACERDA, R.; SILVA, M. Materiais e estratégias didáticas em ambiente virtual de aprendizagem. **Revista Brasileira de Estudos Pedagógicos**, v. 96, n. 243, p. 1694-1708, 2015. doi:10.24109/2176-6681.rbep.96i243.1694.

LUI, C.; HONG, Y.; KIM, J.; *et al.* Open science and multicultural research: Some data, considerations, and recommendations. **Cultural Diversity and Ethnic Minority Psychology**, v. 28, n. 1, p. 1-12, 2022. doi:10.1037/cdp0000541.

MANGHI, P.; *et al.* Digital Libraries. **ACM SIGMOD Record**, 2020. doi:10.1145/3385658.3385669.

MEGHINI, C. Data Preservation. **Data Science Journal**, 2013. doi:10.2481/dsj.grdi-009.

MARTINS, R.; *et al.* Perspectivas de acessibilidade na autoria de cursos e recursos educacionais abertos: um panorama internacional. **SBIE**, 2022. doi:10.5753/sbie.2022.225750.

MATIAS *et al.* Ambiente de compartilhamento de REAs: uma proposta de modelagem, implementação e modelo de negócio. **Revista Semiárido de Visu**, 2016. DOI: 10.31416/rsdv.v4i1.159.

SUMÁRIO

MATUSIAK, K.; *et al.* Finding access and digital preservation solutions for a digitized oral history project. **Digital Library Perspectives**, 2017. doi:10.1108/dlp-07-2016-0025.

MAZZETTI, M.; *et al.* PCTI no Brasil: a relação inovação e sistema produtivo na atual estratégia nacional. **Colóquio - Revista do Desenvolvimento Regional**, v. 17, n. 1, 2019. doi:10.26767/coloquio.v17i1.1581.

MENDES-DA-SILVA, W. Revisão pelos Pares Aberta e Ciência Aberta na Comunidade de Pesquisa em Negócios. **Revista de Administração Contemporânea**, 2019. doi:10.1590/1982-7849rac2019190278.

MORRISON, H. The dramatic growth of open access. **Journal of Interlibrary Loan, Document Delivery & Electronic Reserve**, v. 16, n. 3, p. 95-107, 2006. Disponível em: https://doi.org/10.1300/j474v16n03_11. Acesso em: 2 jan. 2025.

MUCK, L. A produção científica em Acesso Aberto da Odontologia brasileira: uma análise da publicação e do impacto na coleção principal da Web of Science. **Transinformação**, v. 35, n. 1, p. 1-12, 2023. doi:10.1590/2318-0889202335e236642.

NASCIMENTO, A. Periódicos de acesso aberto e taxas de processamento de artigos. **Bibliocanto**, v. 9, n. 2, p. 1-12, 2023. doi:10.21680/2447-7842.2023v9n2id33703.

NASCIMENTO, A.; FELIPE, R. Serviços de suporte à gestão de dados de pesquisa. **Brazilian Journal of Information Science**, v. 15, 2021. doi:10.36311/1981-1640.2021.v15.e02126.

NETTO, R.; *et al.* Análise da matriz curricular de um curso de administração: tecnologias e mercado de trabalho. **Peer Review**, 2023. doi:10.53660/301.prw709.

OLIVEIRA, A.; *et al.* Formação profissional em recursos educacionais abertos. **Revista Semiário de Visu**, v. 3, n. 2, 2015. doi:10.31416/rsdvv3i2.173.

PASQUI, L. Digital curation and long-term digital preservation in libraries. **JLIS it**, 2024. doi:10.36253/jlis.it-567.

PATTERSON, D. Sustainable Digital Preservation Initiatives Benefit from Multi-Pronged Approach. **Evidence Based Library and Information Practice**, 2021. doi:10.18438/ebliip29887.

PAVAN, L.; BARBOSA, M. Financiamento público no Brasil para a publicação de artigos em acesso aberto: alguns apontamentos. **Em Questão**, v. 23, n. 2, p. 120-145, 2017. DOI: 10.19132/1808-5245232.120-145.

SUMÁRIO

PONTIKA, Nancy; KNOTH, Petr; CANCELLERI, Matteo; PEARCE, Samuel. Fostering open science to research using a taxonomy and an elearning portal. In: ACM INTERNATIONAL CONFERENCE ON KNOWLEDGE TECHNOLOGIES AND DATA-DRIVEN BUSINESS (i-KNOW '15), 15., 2015, Graz. Proceedings... New York: Association for Computing Machinery, 2015. p. 1–8. DOI: 10.1145/2809563.2809571.

ROSS-HELLAUER, Tony. What is open peer review? A systematic review. **F1000Research**, London, v. 6, p. 588, 2017. DOI: <https://doi.org/10.12688/f1000research.11369.2>. Disponível em: <https://f1000research.com/articles/6-588>. Acesso em: 29 mar. 2025.

SANTOS, A.; FLORES, M. Infraestrutura de segurança para gestão de riscos do Repositório Arquivístico Digital Confiável: um diálogo com a ISO 16363. **Prisma Com**, 2020. doi:10.21747/16463153/41a5.

SANTOS, M. A.; SANTOS, R. F. Produção de conhecimento em rede: em busca de uma ciência colaborativa, aberta e comum. **Contribuciones a las Ciencias Sociales**, v. 16, n. 7, p. 1-10, 2023. doi:10.55905/revconv.16n.7-079.

SANTOS, R. Protocolos de treinamento físico em modelos experimentais usados na pesquisa em Educação Física. **Colloquium Health and Education**, 2021. doi:10.37497/colloquium.v1i2.16.

SANTOS, R.; MONTEIRO, F. Estudo da percepção de pesquisadores da área de ciências agrárias sobre Acesso Aberto. **Incíd: Revista de Ciência da Informação e Documentação**, v. 4, n. 2, p. 34-53, 2013. doi:10.11606/issn.2178-2075.v4i2p34-53.

SILVA, José Paulo Monteiro da; INÁCIO JÚNIOR, Eudes. Mapeamento da Ciência Aberta em Revistas Brasileiras de Administração: situação atual e perspectivas futuras. In: ABEC MEETING, 2024, São Paulo. Anais [...]. São Paulo: ABEC Brasil, 2024. Disponível em: <https://blog.scielo.org/blog/2025/01/15/a-ciencia-aberta-ainda-e-tema-incipiente-nos-periodicos-de-administracao/>. Acesso em: 29 mar. 2025.

SECO, M.; MENDES, L. OpenStax Connexion versus Wikibooks: Análise comparativa de plataformas de suporte a livros abertos. **Education in the Knowledge Society (EKS)**, 2016. doi:10.14201/eks20161745374.

SHINTAKU, Milton; SALES, Luana Farias (org.). **Ciência aberta para editores científicos**. Botucatu, SP: ABEC, 2019. DOI: 10.21452/978-85-93910-02-9.

SUMÁRIO

SILVEIRA, Lúcia da *et al.* Taxonomia da Ciência Aberta: revisada e ampliada. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 28, p. 1-22, 2023. DOI: 10.5007/1518-2924.2023.e91712. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/91712>. Acesso em: 3 jan. 2025.

SOUZA, L. G. S.; AGANETTE, E. C. Preservação digital em longo prazo amparada por planos de ações. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 18, 2020. Disponível em: <https://doi.org/10.20396/rdbci.v18i0.8661185>. Acesso em: 3 jan. 2025.

VASCONCELOS, F. Inovação aberta em companhias de saneamento básico brasileiras. **Read: Revista Eletrônica de Administração**, Porto Alegre, 2023. doi:10.1590/1413-2311.393.122183.

XU, F.; JUTY, N.; GOBLE, C.; JUPP, S.; PARKINSON, H.; COURTOT, M. Features of a fair vocabulary. 2022. Disponível em: <https://doi.org/10.21203/rs.3.rs-1465079/v1>. Acesso em: 2 jan. 2025.

ZANIN, G. Recursos educacionais abertos e direitos autorais: análise de sítios educacionais brasileiros. **Revista Brasileira de Educação**, 2017. doi:10.1590/s1413-24782017227174.



11

Rafael Port da Rocha

**DIGITALIZAÇÃO
E PRESERVAÇÃO
DIGITAL:
ABORDAGENS METODOLÓGICAS**

DOI: 10.31560/pimentacultural/978-85-7221-475-9.11

SUMÁRIO

COMO CITAR:

ROCHA, Rafael Port da. Digitalização e preservação digital: abordagens metodológicas. /n: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação: Fundamentos e Aplicações**. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

A preservação digital peneja e realiza ações gerenciais e técnicas para garantir que os objetos digitais sejam acessados e usados pelo tempo que for necessário, por meio dos recursos computacionais disponíveis no momento que esse uso ocorrer, ainda que esse momento ocorra em um futuro distante.

Com o passar dos tempos é previsível que ocorram mudanças no ambiente de um objeto digital, como em software ou hardware, e estas mudanças podem colocar em risco a preservação deste objeto. Por exemplo, a obsolescência de um formato de um arquivo é uma mudança decorrente da evolução do ambiente tecnológico. Essa mudança pode colocar em risco a preservação de objetos digitais que estão representados nesse formato, pois os ambientes computacionais poderão não dispor mais de recursos para manipulação de objetos que estão nesse formato obsoleto.

O planejamento da preservação digital envolve identificar situações em que mudanças no ambiente colocam em risco preservação de um objeto digital, e também especificar estratégias e ações para mitigar esses riscos. Por exemplo, a ação de migração de formato é especificada para mitigar o risco da obsolescência de formato de arquivo. Nesse caso, uma nova representação para o objeto digital é gerada em um formato de arquivo atual. Outro risco

SUMÁRIO

de preservação decorre da obsolescência dos dispositivos de armazenamento (como discos rígidos ou discos óticos). Nesse caso, a ação de migração de mídia é planejada, através da qual os objetos digitais são transferidos de uma mídia em tecnologia obsoleta para uma mídia tecnologicamente atual.

A preservação digital também envolve mitigar riscos decorrentes de deterioração, falha ou desastre que possam ocorrer com os dispositivos que armazenam os objetos digitais. Para tal, ações de preservação digital preveem, por exemplo, a realização de cópias de segurança, que devem ser armazenadas em locais distantes. Para mitigar riscos de corrupção dos arquivos que representam objetos digitais, ações de verificação periódica da integridade da informação armazenada são planejadas.

Um objeto digital não está exposto somente a riscos decorrentes de mudanças do ambiente tecnológico. Muitos acervos digitais deixam de existir por falta de previsão continuada de recursos financeiros, por falta de recursos humanos ou de um programa de qualificação continuada para a equipe. Acervos digitais também deixam de existir devido a riscos como pouco comprometimento institucional com os acervos, fraca responsabilização pela preservação dos objetos digitais, falta de segurança nas instalações físicas, entre outros.

No cenário atual, os avanços das pesquisas em preservação digital estão sendo incorporados em melhores práticas para gestão de repositórios de objetos digitais, apoiadas por normas, recomendações e mecanismos de certificação. Em especial, destacam-se o modelo de referência para repositório digital OAIS (The Reference Model for an Open Archival Information System) (CCSDS , 2024a), que é norma internacional ISO 14721 e norma brasileira ABNT NBR ISO 14721, e os critérios para certificação de repositórios digitais confiáveis CoreTrustSeal (Core Trust Seal, 2022) e ACTDR (Audit and Certification of Trustworthy digital Repositories) (CCSDS , 2024b), que é a norma internacional ISO 16363.

SUMÁRIO

A digitalização de acervos está atrelada à preservação digital, à medida que carece observar a motivação principal da preservação digital: os objetos digitais digitalizados devem ser usados pelo tempo que for necessário, por meio dos recursos computacionais disponíveis no momento que esse uso ocorrer. Isso implica que a digitalização de um acervo necessita compartilhar e aplicar aspectos metodológicos, técnicas, estratégias e ações de preservação digital.

Este capítulo aborda aspectos metodológicos da digitalização de um acervo na perspectiva de que este acervo digital será preservado pelo tempo que for necessário. Por isso, apresenta aspectos metodológicos da digitalização e sua inter-relação com aspectos metodológicos da preservação digital.

O capítulo está estruturado da seguinte forma: Apresenta o que é um objeto digital, na perspectiva de que este deve ser capaz de representar um material digitalizado (seção 1). Identifica que objetos digitalizados possuem características que não devem ser perdidas ao longo de transformações decorrentes de ações de preservação digital, como migração (seção 2). Relaciona que objetos digitalizados são descritos por meio de metadados descritivos, estruturais, técnicos e de preservação digital (seção 3) e empacotados (seção 4). Identifica que mudanças no ambiente podem colocar a preservação de um objeto digitalizado em risco e que ações para mitigar esses riscos devem ser planejadas (seção 5). Apresenta as matrizes e as derivadas de acesso geradas pela digitalização na perspectiva que são representações de um objeto digital (seção 6). Apresenta aspectos metodológicos da digitalização contextualizados com preservação digital (seção 7). Observa a necessidade em preservar em longo prazo os materiais digitalizados e traz aspectos metodológicos de preservação digital envolvidos (seção 8). Apresenta uma estratégia em que as abordagens metodológicas para digitalização e preservação digital são apresentadas de por meio de corpos de conhecimento. (seção 9). Nessa estratégia, corpos de conhecimento em digitalização e curadoria digital trazem, de forma integrada, conhecimentos e

SUMÁRIO

práticas em digitalização e preservação digital, cujos recursos envolvidos (como objeto digital, processos, documentos e diretivas) são descritos por ontologia. Termos relacionados com aspectos metodológicos em preservação e digitalização, que são mencionados e relacionados em várias análises, são enfatizados em **negrito**.

O OBJETO DIGITAL E SUAS REPRESENTAÇÕES

Na perspectiva da preservação digital, um objeto digital é uma unidade discreta de informação sujeita à preservação digital. Um objeto digital é uma entidade complexa que envolve vários níveis de abstração (intelectual, lógico e físico) e é visto como uma entidade intelectual, que possui uma ou mais representações digitais. Essas representações digitais são formadas por arquivos digitais.

No nível intelectual, um objeto digital é percebido como uma criação intelectual ou artística e é chamado de entidade intelectual. Nesse nível, o objeto digital assume formas ou concepções familiares aos seres humanos. Os objetos digitais são vistos sob a ótica dos seus propósitos de uso pelas pessoas, isto é, como artigos, atas, monografias, notas fiscais, fotografias.

No nível intelectual, a maneira com que informação é concretizada não é relevante, isto é, não é relevante se objeto digital é concretizado por meio de páginas representadas por imagens em formato TIFF e armazenadas em um disco rígido, ou se o objeto digital é representado por meio de um arquivo digital, no formato PDF, em um disco ótico. A percepção do objeto digital como uma entidade intelectual está ligada à configuração da informação, que nos permite identificar que um objeto digital é um artigo, um livro, uma foto, um vídeo. Nesse nível, metadados descritivos são usados para descrever os objetos digitais para propósitos como descoberta e identificação.

SUMÁRIO

Um objeto digital é visto em nível lógico na perspectiva que é representado por meio de pastas e arquivos, e que estes são constituídos por sequências de números binários (bits) que seguem regras e padrões de interpretação. Nesse contexto, os formatos de arquivo indicam as regras que indicam como os bits que constituem os arquivos digitais são codificados.

No nível lógico, uma entidade intelectual pode assumir várias representações digitais. Por exemplo, uma obra rara digitalizada pode ser representada por meio de um único arquivo PDF ou pode ser representada por múltiplos arquivos de imagem, com cada arquivo um sendo a imagem de uma página, no formato TIFF.

Metadados estruturais são usados para indicar como elementos que compõem um objeto digital são estruturados. Esses metadados descrevem a estrutura do objeto digital (como seus capítulos e seções), suas páginas e os arquivos que contém o objeto e os metadados. Por exemplo, no caso de uma obra rara digitalizada sendo representada por páginas individualizadas em arquivos no formato TIFF, metadados estruturais descrevem as seções e as páginas da obra rara, e mapeiam essas seções e páginas para os respectivos arquivos.

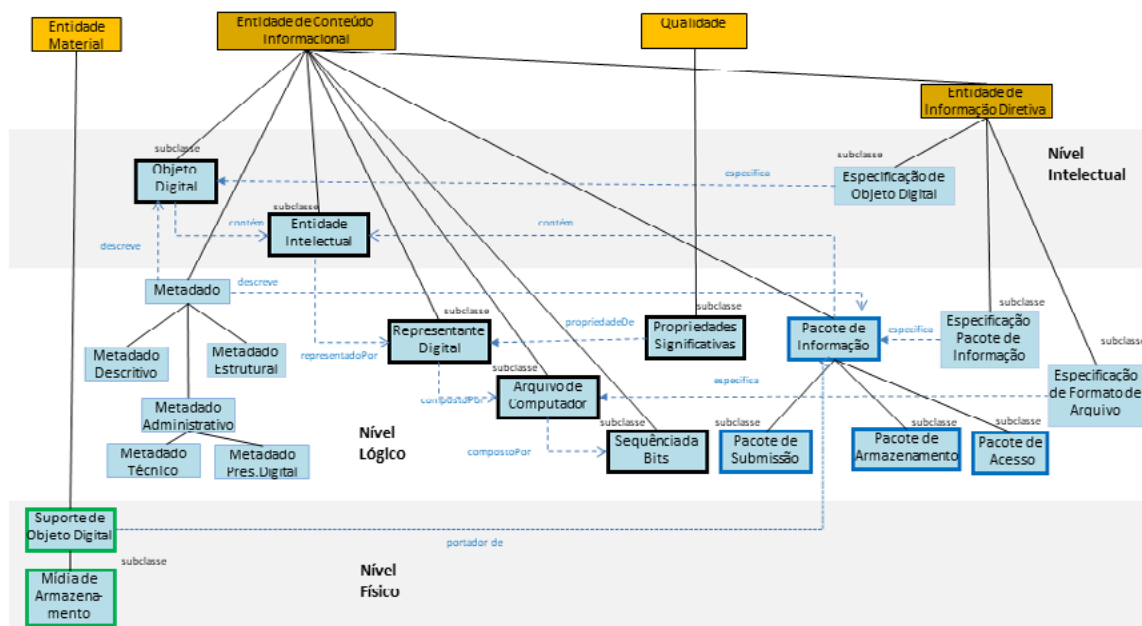
Metadados técnicos descrevem atributos e propriedades das representações dos objetos digitais, como formato de arquivo, resolução de imagem, método de compressão de imagem, software e hardware necessários, informações para checagem de integridade (*checksum*) e para garantir autenticidade (assinatura digital).

O objeto digital em nível físico compreende no armazenamento das representações em um suporte, como um disco ótico, uma fita magnética ou disco rígido. É importante destacar que um objeto digital não existe sem ter suas representações digitais afixadas em um suporte.

SUMÁRIO

Abordagens metodológicas que abordam objeto digital são apresentadas em Dappert e Enders (2010), PREMIS Editorial Committee (2015) e Pires e Rocha (2020). A figura 1 apresenta o objeto digital, sua entidade intelectual, representações, arquivos, e suporte.

Figura 1 - objeto digital, Pacote de Informação e Metadados



AS PROPRIEDADES SIGNIFICATIVAS DE UM OBJETO DIGITAL

A relação entre a entidade intelectual de um objeto digital e as suas representações envolve garantir que as representações deste objeto digital devem atender às expectativas de valor atribuídas ao objeto digital por sua comunidade de usuários, na perspectiva que este é visto como uma entidade intelectual. O curador digital deve compreender essa percepção de valor e estabelecer critérios para que não sejam criadas Representações sem valor para estes usuários.

SUMÁRIO

Esses critérios são chamados de características significativas, que podem variar para cada objeto digital. Por exemplo, o tipo de letra pode não ser uma característica significativa para as atas de um departamento, mas pode ser imprescindível para os usuários de um jornal, na percepção de que os aspectos tipográficos são considerados parte da natureza desse jornal. Isso significa que, em ações futuras de migração de formato, as novas representações do jornal devem manter a mesma fonte de letra. Já para as atas do departamento, se estas forem migradas para documentos com textos representados em outra fonte, continuarão mantendo seu valor como ata para o departamento.

O planejamento da preservação envolve identificar e registrar no plano de preservação digital quais são as propriedades significativas de um objeto digital. Dessa forma, quando houver a execução de uma ação de preservação digital de migração de formato, esta deverá ser obedecer à restrição das propriedades significativas, na geração das novas representações.

Estudos em propriedades significativas na perspectiva na perspectiva das entidades intelectuais são abordadas em Dappert e Farqhar (2009a), Dappert (2013) e Pires e Rocha (2020). A figura 1 apresenta o objeto digital e suas propriedades significativas.

OS METADADOS DO OBJETO DIGITAL

Metadados são classificados em descritivos, estruturais e administrativos, sendo que os metadados administrativos são subdivididos em metadados técnicos e de preservação digital.

Os metadados descritivos descrevem o objeto com foco que este é uma entidade intelectual. Para tal, são usados esquemas padrões de metadados adequados para a natureza dos objetos digitais.

SUMÁRIO

Por exemplo, o MARC é usado para descrever materiais bibliográficos, ISAD-G é usado para descrever acervos de natureza arquivística, DDI e DataCite são usados para descrever dados de pesquisa, e Dublin Core é uma linguagem comum e simples para descrever qualquer tipo de recurso da Web. Em muitos casos, objetos digitais são descritos por mais de um esquema padrão de metadados. Por exemplo, acervos arquivísticos trazem a descrição dos seus itens em ISAD-G e em Dublin Core. Repositórios de dados de pesquisa normalmente usam os esquemas DDI, DataCite e Dublin Core.

Metadados estruturais descrevem os componentes do objeto digital, isto é, suas representações, os arquivos que compõe essas representações e os arquivos que registram os metadados. Metadados estruturais também permitem a descrição da estrutura da informação do objeto digital (como páginas, seções, capa, sumário, índice remissivo) e mapeiam elementos dessa estrutura para arquivos correspondentes nas representações. METS e o esquema padrão usado para metadados estruturais.

Metadados técnicos descrevem aspectos técnicos de um objeto digital. Descrevem os formatos dos arquivos que fazem parte da representação de um objeto digital. Também descrevem aspectos como resolução e profundidade e o modo, no caso de imagem, tipo de letra, no caso de documento textual, entre outros. Além disso, são considerados Metadados técnicos aqueles que servem para evidenciar a integridade e autenticidade do objeto digital, como *checksum* e assinatura digital de um arquivo.

A produção dos metadados técnicos é normalmente realizada de forma automatizada por meio de ferramentas como JHOVE, FITS, ExifTool e FFprobe. MIX e Exif são esquemas padrões para metadados técnicos voltados para fotografias e imagens.

Metadados técnicos são usados no planejamento e na execução de ações de migração, para selecionar os objetos digitais que estão em risco, pois são representados por arquivos em formato obsoleto.

SUMÁRIO

metadados técnicos que registram *checksum* são usados em ações de preservação digital para verificar a integridade dos arquivos, verificando se o *checksum* de um arquivo que foi registrado em metadados confere com o checksum do arquivo atualmente armazenado.

Os metadados de preservação digital são metadados administrativos que tratam da proveniência de um objeto digital e da gestão de seu arquivamento. Informações de proveniência são informações sobre entidades, atividades e pessoas envolvidas na produção de um objeto digital, que podem ser usadas para formar avaliações sobre sua qualidade ou confiabilidade (Moreau et al, 2013). Descrevem, por exemplo, ações (e agentes envolvidos) realizadas em um objeto digital, como sua ingestão no repositório digital, verificação de sua integridade (*checksum*), migração de formato, entre outras. PREMIS (PREMIS Editorial Committee, 2015) é o esquema padrão para metadados de preservação digital.

Dappert e Enders (2010) apresentam os padrões de metadados sob a ótica da preservação digital. A figura 1 apresenta os metadados descritivos, estruturais, técnicos e de preservação digital, que descrevem objetos e pacotes de informação.

O PACOTE DO OBJETO DIGITAL

Um objeto digital é armazenado por meio de Pacote de Informação, que é o recipiente deste objeto digital. Um pacote de informação é composto:

- Pelas pastas e arquivos das representações do objeto digital.
- Pelos arquivos que contém os metadados descritivos, técnicos e de preservação digital.

SUMÁRIO

- Pelos arquivos que contém metadados estruturais, que descrevem os elementos que fazem parte do objeto digital, como sua estrutura (capítulos, seções), suas representações e seus arquivos, e arquivos que registram os metadados.

Na perspectiva da digitalização, armazenamento e disseminação de um objeto digital, o modelo OAIS identifica três tipos de pacote de informação: o pacote de informação de submissão, que contém o objeto digitalizado, que é enviado para ser armazenando em um repositório; o pacote de armazenamento de informação, que contém o objeto digital recebido da digitalização, verificado e armazenado pelo repositório; e o Pacote de Disseminação de Informação, que é o objeto digital entregue ao usuário, e contém o objeto na representação adequada para o seu uso.

E-ARK CSIP (DILCIS Board, 2021) é uma especificação europeia para os pacotes de submissão, armazenamento e disseminação estabelecidos pelo modelo OAIS, com estruturas descritas pelo padrão METS e indicações do uso de metadados de preservação digital (PREMIS) e descritivos. BagIt (Kunze *et al.*, 2018) é uma especificação para armazenamento e transferência de conteúdos digitais. Estabelece convenções hierárquicas de layout dos arquivos, incluindo manifesto de carga, que fornece uma listagem dos arquivos que foram carregados no “bag” e seus checksums, e arquivos de “tags”, que contém metadados sobre o bag. A figura 1 apresenta os pacotes de submissão, armazenamento e disseminação da informação, que contém objetos digitais e são descritos por metadados.

SUMÁRIO

O AMBIENTE, O RISCO E O PLANEJAMENTO DA PRESERVAÇÃO

Uma ação de preservação de um objeto digital é necessária quando a preservação deste objeto está em risco. Esta situação de risco ocorre porque houve mudanças no ambiente do objeto digital.

Esse ambiente pode ser tanto tecnológico, quanto administrativo. Por exemplo, a obsolescência de um formato está ligada ao ambiente tecnológico, pois coloca em risco o uso de objetos em formatos obsoletos em uma plataforma atual. Já o risco da descon continuidade do acervo por falta de recursos está ligado ao ambiente administrativo, que não criou mecanismos institucionais suficientes para garantir o financiamento continuado para armazenar e preservar os objetos digitais.

Para a preservação digital é fundamental compreender e monitorar as mudanças que irão ocorrer nos ambiente tecnológico e administrativo do objeto digital. É a partir desse monitoramento que os riscos de preservação serão percebidos e ações de preservação serão realizadas para mitigar esses riscos.

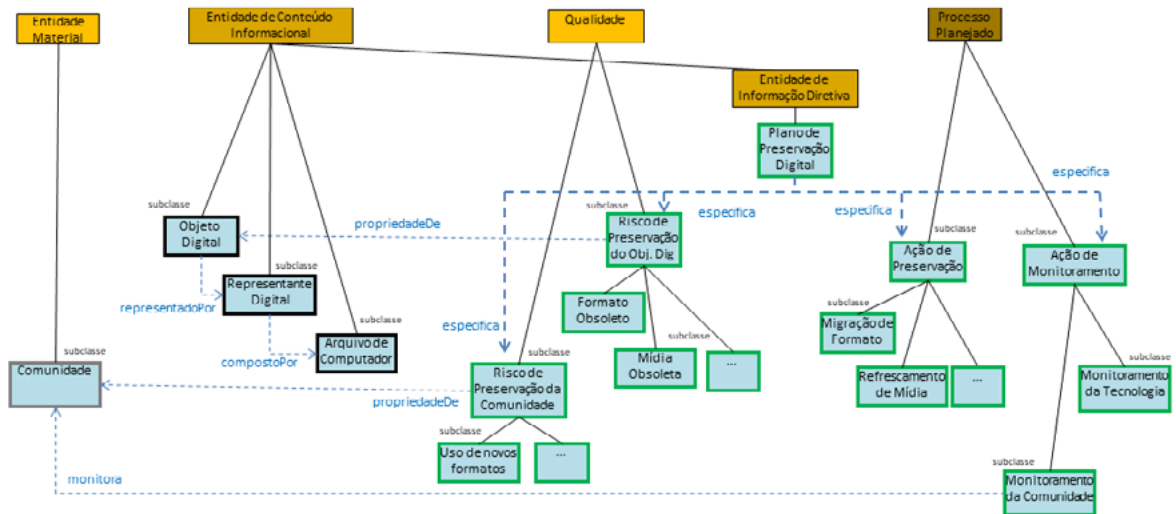
Faz parte do mandato da preservação, planejar e executar a preservação. O plano de preservação estabelece ações de preservação digital que serão realizadas para mitigar riscos de preservação. Estabelece que a preservação digital deverá monitorar os ambientes tecnológico e gerencial, objetivando identificar riscos e ativar ações de preservação adequadas para mitigar esses riscos. O plano de preservação também estabelece as propriedades significativas dos objetos digitais, indicando características dos objetos digitais (como tipo de letra, margem, cor) que deverão ser observadas na geração de representações digitais que são resultados de ações de preservação, como, por exemplo, uma nova representação digital de um objeto cujo formato de arquivo passou a ser obsoleto.

SUMÁRIO

No sentido de mitigar o risco da organização não assumir sua responsabilidade pela preservação dos objetos digitais, é fundamental que esta estabeleça sua política de preservação digital, e que esta política seja conhecida e aprovada institucionalmente. Uma política de preservação é uma declaração formal em que a instituição assume a responsabilidade pela preservação e estabelece diretrizes e orientações de como a organização irá realizar seu mandato de preservação digital.

Abordagens metodológicas que abordam o ambiente e o riscos relacionados à preservação digital são apresentadas por Dappert (2013) e (Dapper e Farqhar, 2009 b). A figura 2 apresenta o plano de preservação digital, que especifica riscos de preservação digital, ações de monitoramento de mudanças na comunidade e na tecnologia e ações de preservação que mitigam riscos de preservação digital.

Figura 2 - Risco, Ações e Planejamento da Preservação Digital



O OBJETO DIGITAL NA DIGITALIZAÇÃO DE UM ACERVO

Um objeto a ser digitalizado resultará em um objeto digital, que é uma entidade intelectual, composta por representações digitais, estas formadas por arquivos. A percepção de que um objeto a ser digitalizado é uma entidade intelectual demanda por um tratamento especializado de curadoria que envolve a caracterização do objeto original como uma criação ou obra intelectual. Isso é determinante na escolha do esquema de metadados descritivos a ser usado para descrever o objeto, assim como para definir as propriedades significativas deste objeto, isto é, as propriedades que o objeto digital não poderá perder quando for digitalizado e preservado em longo prazo.

A digitalização é concretizada quando representações digitais são geradas a partir do objeto original por meio de equipamentos de captura (scanners). Um objeto digitalizado pode resultar em várias representações digitais, cada uma sendo adequada para um uso distinto. Uma representação chamada matriz é gerada em qualidade mais próxima ao original e será administrada pela preservação digital. Serão monitoradas mudanças que colocam a preservação dessa matriz em risco, e ações de preservação serão ativadas para mitigar esses riscos. Representações chamadas de derivadas se acesso são geradas para serem disponibilizadas aos usuários, em formatos adequados às características dos dispositivos e/ou características especiais de acesso. As derivadas de acesso são mais leves que as matrizes, são geradas a partir das matrizes e novas derivadas de acesso podem ser geradas para atender à novas demandas de acesso.

Por exemplo, na digitalização de uma obra rara, sua representação matriz poderia ser composta por arquivos TIFF, cada um representando uma página, com imagens não compactadas em alta resolução. Já uma representação a ser usada para trocas por meio

SUMÁRIO

de mensagens eletrônicas poderia ser representada por um único arquivo no formato PDF, com o uso de técnicas de OCR para tornar o texto pesquisável. Já, para ser acessada na web, essa obra rara poderia ter uma representação com cada página sendo um arquivo no formato JPEG, compactado com perda, sendo acessada através de um visualizador de páginas.

A DIGITALIZAÇÃO NO CONTEXTO DA PRESERVAÇÃO DIGITAL

A digitalização de um acervo envolve: seleção, conversão, controle de qualidade, descrição e representação, produção e gerenciamento da digitalização. Essas capacidades devem estar alinhadas com a preservação digital.

Na seleção são verificados os propósitos, as comunidades alvos, os benefícios, as capacidades e os riscos da digitalização, no sentido de dar andamento a um empreendimento que tenha sucesso. A digitalização de um acervo amplia o acesso ao material, preserva a manipulação dos originais, assim como proporcionam redundância necessária para mitigar riscos de perdas dos originais.

Na seleção são verificados possíveis impedimentos legais (como direito autoral), se os atributos físicos dos originais permitem o planejamento de uma digitalização com riscos mínimos aceitáveis aos originais, se o material já não foi digitalizado, se a instituição tem capacidades para gerar e manter o material digital; assim como verificar os custos e a disponibilidade dos recursos financeiros, tecnológicos e humanos necessários para o empreendimento. É importante que a digitalização só ocorra após o acervo ter recebido o tratamento técnico especializado, respeitando sua característica bibliográfica, arquivística ou museal. É um grande risco digitalizar um acervo que

SUMÁRIO

não foi tecnicamente tratado, pois, dessa forma, o mesmo não será de uso adequado para sua comunidade de usuários. O tratamento técnico normalmente já envolve a produção dos metadados descritivos.

A seleção aborda aspectos que impactam na preservação digital. Cabe à seleção demonstrar que o acervo atende a uma comunidade de usuários (comunidade alvo), e isso implica que seleção deve observar os objetos desse acervo como produções intelectuais ou artísticas, isto é, como entidades intelectuais, incluindo, consequentemente, a discussão em torno das propriedades significativas desses objetos. A seleção ainda avalia riscos que também são riscos no contexto da preservação digital, como riscos de impedimentos por aspectos legais, falta de recursos ou infraestrutura adequada.

A seleção deve analisar o acervo a ser digitalizado em uma perspectiva que vai além da entrega do material digitalizado. Não incluir a preservação digital na seleção é um fator que pode impactar negativamente no sucesso do empreendimento, pois sem a preservação digital todos os esforços e benefícios da digitalização poderiam ser anulados com uma eventual perda futura do acervo.

Demonstrada a viabilidade do empreendimento de digitalização, a conversão segue a seleção. A conversão envolve a definição das configurações para os objetos que serão digitalizados, a partir da análise de atributos físicos dos originais e de suas características de produção e de conservação. A partir dessa análise, são definidas as configurações para os itens a serem digitalizados. O estudo realizado pela conversão também irá subsidiar a escolha dos equipamentos de captura mais adequados, como scanners planetários, para o caso de livros raros e encadernados ou scanners de grandes formatos para mapas.

Por exemplo, no caso de digitalização de documentos textuais com ilustrações, na conversão são avaliados aspectos técnicos como dimensão, níveis de detalhes e presença de cor, entre outros.

SUMÁRIO

São consideradas características de produção ou conservação, como o material ser um texto impresso, ou ser texto em traço, assim como ter a presença de imagens em meio tom, presença de manchas, etc. A partir dessa análise, são definidas configurações para os itens a serem digitalizados, como profundidade de bit, cor, compressão, formato de arquivo, entre outros.

Em alinhamento com a preservação digital, a conversão é o estudo que vai subsidiar a especificação das representações dos objetos digitais, na perspectiva que estes objetos digitais já foram investigados na seleção como entidades intelectuais. A conversão envolve identificar as configurações para as representações digitais que são a matriz e as derivadas de acesso. É importante considerar que as configurações escolhidas para essas representações devem atender o que foi especificado como propriedades significativas.

Com base nas configurações estabelecidas pela conversão, a descrição e a representação envolvem, respectivamente, identificar os esquemas padrões de metadados a serem usados e especificar como os objetos serão digitalmente representados.

Em alinhamento com a preservação digital, descrição e representação envolve especificar as representações do objeto digital tendo como requisitos parâmetros estabelecidos na configuração. Envolve também identificar os esquemas para os metadados descritivos, estruturais, técnicos e de preservação digital, que também fazem parte dessas representações. Resulta na especificação das representações dos objetos digitais, incluindo suas pastas, arquivos, nomes de arquivos e de pastas, formatos de arquivos, esquemas de metadados e arquivos de metadados. Também resulta na especificação dos pacotes de submissão, armazenamento e disseminação, que são os recipientes destes objetos digitais. O pacote de submissão é o resultado da digitalização, que, depois de ingerido pelo repositório, é arquivado como pacote de armazenamento e preservado em longo prazo. Esse pacote de armazenamento representa a matriz. Já os pacotes de acesso estão relacionados às derivadas de acesso.

SUMÁRIO

O controle de qualidade tem como objetivo garantir que as expectativas de qualidade sejam cumpridas. Para tal, são estabelecidos critérios que servem pra avaliar se o material foi digitalizado corretamente, como dimensão, resolução, nome de arquivo, formato de arquivo modo, cor, recorte, rotação, inclinação, nitidez da imagem, sem presença de artefatos digitais, entre outros. Esses critérios são usados em uma rotina de controle de qualidade, que deve fazer parte do processo de digitalização, em que materiais que não atendem aos critérios de qualidade não deverão ser aceitos.

Na perspectiva da preservação digital, critérios de qualidade ampliam a especificação das representações para os objetos digitais. A realização da rotina de controle de qualidade contribui no sentido de prover evidências de autenticidade aos objetos digitais em relação aos seus originais. Essas evidências só irão existir se os critérios forem documentados e a realização das rotinas de controle de qualidade for registrada por meio de metadados de preservação digital. Na preservação digital, o controle de qualidade visa garantir que os objetos digitais sejam criados e preservados em conformidade com suas especificações.

A produção envolve planejar e executar o fluxo de digitalização. Esse fluxo é composto por atividades como: receber os originais, preparar o material para digitalização, digitalizar o material, obter os metadados descritivos (pois o material já foi previamente tratado e descrito), compor os representantes dos objetos digitais, gerar e empacotar os objetos digitais, controlar a qualidade, devolver os originais e entregar os pacote dos objetos digitais aos ambientes de armazenamento em longo prazo e acesso.

Na perspectiva do modelo de referência para repositórios digitais OAIS, os ambientes de armazenamento em longo prazo e acesso são especificados através do modelo funcional, que é composto pelas entidades funcionais de ingestão, armazenamento, planejamento da preservação, acesso, administração e gerenciamento

SUMÁRIO

de dados. A ingestão recebe o pacote de submissão de informação resultante da digitalização, verifica e garante que o mesmo está de acordo com as especificações e o entrega ao armazenamento como pacote de armazenamento da informação. O acesso torna os objetos digitais e serviços de acesso (entrega, busca) visíveis a seus consumidores. São entregues aos consumidores os pacotes de disseminação, que envolvem derivadas de acesso adequadas a esses consumidores.

Proveniência e autenticidade são requisitos fundamentais para a preservação dos objetos digitais. No sentido de prover evidências de proveniência e autenticidade a estes objetos, é determinante que a produção dos mesmos, isto é a digitalização, seja especificada, e que as atividades da digitalização sejam realizadas de acordo com essa especificação, assim como registradas por meio de metadados de preservação, incluindo os responsáveis pelas realizações das atividades.

Abordagens metodológicas que orientam a digitalização são apresentadas em KENNEDY *et al.* (2003), Nascimento *et al.* (2006), Fiocruz (2019) e CONARQ (2010). A figura 3 apresenta em tabela processos relacionadas com seleção, conversão, descrição e representação, controle de qualidade e produção.

PRESERVANDO OS MATERIAIS DIGITALIZADOS

Encerrada uma digitalização, a missão seguinte é armazenar, preservar e dar acesso aos materiais digitalizados pelo tempo que for necessário. Isso é especificado pelo modelo OAIS e realizado por repositórios digitais que são planejados de acordo com OAIS.

SUMÁRIO

O modelo OAIS é composto pelo modelo de informação, que especificam os objetos digitais e os pacotes de informação, e pelo modelo funcional, que especifica as funções devem ser realizadas para cumprir a missão de armazenar, preservar e dar acesso.

Em relação ao modelo funcional de OAIS, além da ingestão e do Acesso já trazidas na seção anterior como destinos da digitalização, o armazenamento envolve receber, armazenar, entregar e manter a integridade (via *checksum*) dos pacotes de armazenamento da informação. Também cabe ao armazenamento gerenciar as mídias usadas e repor e recuperar mídias após desastre.

O planejamento da preservação envolve desenvolver o plano de preservação digital. Esse plano especifica os riscos (como obsolescência de formato), as ações de preservação que serão ativadas para mitigar esses riscos (como migração de formato), assim como as propriedades significativas. Também faz parte do planejamento da preservação digital, documentar as especificações dos objetos digitais e seus pacotes. Além disso, planejamento da preservação fornece ações de monitoramento da comunidade e da tecnologia, que objetivam, respectivamente, identificar mudanças na comunidade e na tecnologia que podem colocar em risco a preservação dos objetos digitais, e ativar ações planejadas de preservação digital para mitigar esses riscos.

A administração controla a operação das outras entidades funcionais, envolvendo, entre outras atividades, estabelecer e manter padrões, normas e políticas; negociar e auditar submissões; gerenciar a configuração de sistemas de software e hardware. Já a gestão de dados realiza o provimento, a manutenção e o acesso das informações descritivas sobre os objetos digitais e sobre os documentos e dados usados para gerenciar ambiente especificado por OAIS. Também provém consultas e relatórios envolvendo essas informações descritivas.

SUMÁRIO

No sentido de certificar que repositórios operam em conformidade com OAIS e, conseqüentemente, preservam adequadamente seus conteúdos pelo tempo que for necessário, instrumentos de certificação como TRAC, ACTDR e CoreTrustSeal foram desenvolvidos. Esses instrumentos estabelecem critérios de certificação, que são organizados em três categorias: infraestrutura organizacional, gestão do objeto digital e infraestrutura tecnológica e segurança da informação.

Os critérios relativos à infraestrutura organizacional estão relacionados a riscos referentes a deficiências em aspectos como: reconhecimento e comprometimento institucional com a preservação digital, governança, viabilidade organizacional, responsabilidades, fluxos e rastreabilidade, recursos financeiros e humanos, qualificação da equipe, gestão de direitos, aspectos éticos, privacidade da informação, contratos e licenças. Os critérios relativos à gestão do objeto digital estão relacionados a riscos ligados aos objetos digitais e estão focados em certificar da realização das entidades funcionais do modelo OAIS, já que as mesmas estabelecem as atividades a serem realizadas para manter esses objetos, incluindo o planejamento da preservação. Já os critérios referentes à infraestrutura tecnológica envolvem ambiente de software e hardware, integridade do armazenamento (cópias de segurança, detecção de corrupção ou perda de dados e sua recuperação), e critérios de segurança da informação envolvem riscos de segurança associados aos dados, sistemas, equipe e instalações físicas.

É importante destacar que na perspectiva de OAIS um repositório digital não se resume um serviço informatizado pelo uso de um software de repositório, como DSpace ou Dataverse. Um repositório em conformidade com OAIS é aquele que possui, realiza e certifica sua infraestrutura organizacional, gestão dos objetos digitais, infraestrutura tecnológica e segurança da informação. Softwares e serviços de TI fazem parte desse complexo. Abordagens metodológicas que orientam preservação são apresentadas em CCSDS (2024a), Core Trust Seal (2022), CCSDS (2024b) e Pires e Rocha (2020).

CORPOS DE CONHECIMENTOS EM DIGITALIZAÇÃO E PRESERVAÇÃO DIGITAL

Um corpo de conhecimento representa a agregação de conhecimento de uma área particular que é esperada ser aprendida a fundo por um indivíduo para ser considerado ou certificado como um praticante (Ören, 2005). No sentido de apoiar o desenvolvimento da digitalização alinhada com a preservação digital, os corpos de conhecimento em digitalização¹ e curadoria digital² foram desenvolvidos por Rocha (2025b e 2025c). Estes corpos de conhecimento trazem, de forma integrada, conhecimentos e práticas em digitalização e preservação digital.

Cada corpo de conhecimento é composto por áreas de conhecimento. Cada área de conhecimento corresponde a uma delimitação de conhecimento, e as atividades práticas de cada área de conhecimento são representadas por processos. Um processo possui entradas, saídas e recursos necessários para produzir as saídas a partir das entradas. O corpo de conhecimento também contém informações diretivas, que são aquelas que especificam processos e documentos. Todos os elementos que compõem os corpos de conhecimento, como corpos de conhecimento, áreas de conhecimento, processos, documentos, informações diretivas e outros recursos usados, são especificados pela Ontologia de Corpo de Conhecimento³ (Rocha, 20205a). Essa ontologia usa Basic Formal Ontology (BFO) (Arp; Smith e Spear, 2015) como ontologia de topo. BFO é uma ontologia em que a realidade é representada por continuantes e ocorrentes. Continuantes aquelas entidades que persistem ao longo do tempo, enquanto mantém sua identidade,

1 https://www.ufrgs.br/fabrico/bok/index.php?title=Corpo_de_Conhecimento_em_Digitalização

2 https://www.ufrgs.br/fabrico/bok/index.php?title=Corpo_de_Conhecimento_em_Curadoria_Digital

3 Ontologia Corpo de Conhecimento - https://www.ufrgs.br/fabrico/bok/index.php?title=A_Ontologia

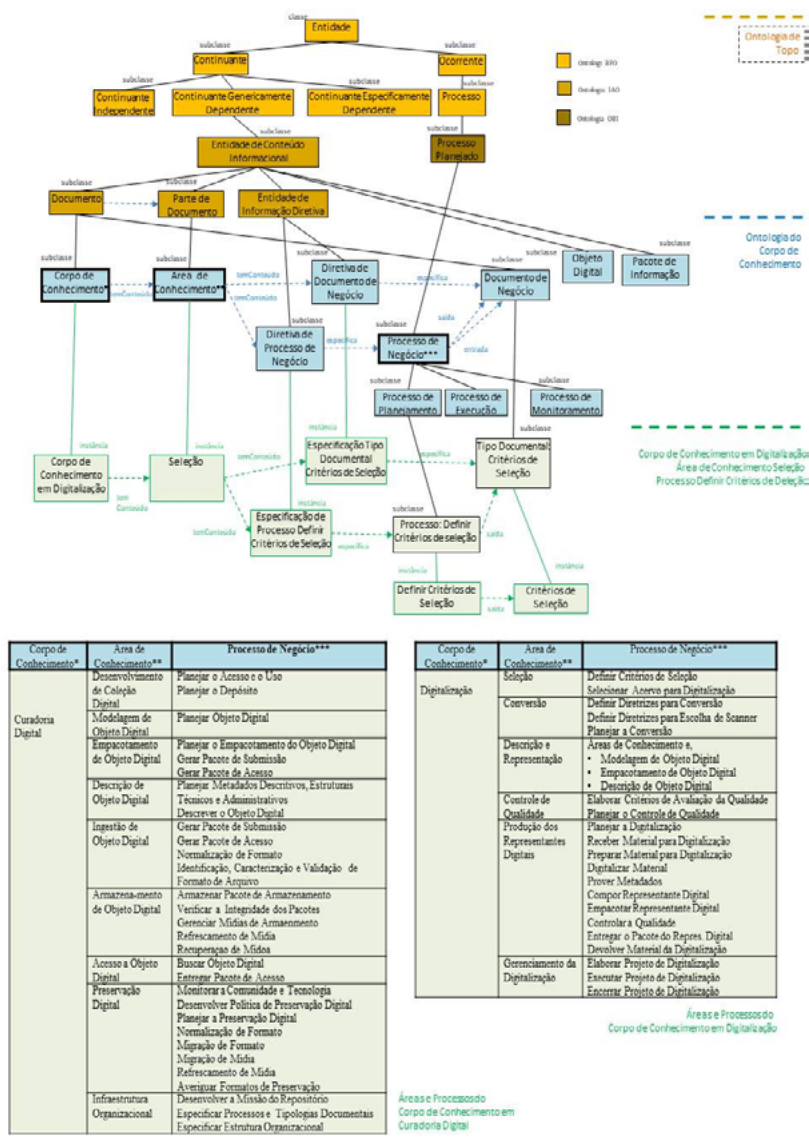
SUMÁRIO

como pessoas e documentos. Já ocorrentes são aquelas se desenrolam ao longo, podendo ter partes temporais, como processos. O corpo de conhecimento está representado em Semantic Media Wiki⁴, que é o ambiente wiki utilizado pela Wikipedia ampliado com recursos para representação de ontologias.

A figura 3 apresenta classes da Ontologia do Corpo de Conhecimento que representam um corpo de conhecimento, suas áreas de conhecimento, os processos de negócio dessas áreas, assim como documentos que são entradas e saída de processos. A figura 3 também exemplifica parte do corpo de conhecimento em digitalização: o corpo próprio corpo de conhecimento, sua área de conhecimento de seleção, e o processo que define critérios de seleção, e sua especificação. As tabelas apresentadas na figura 3 mostram, de forma sintética e parcial, as áreas de conhecimento dos corpos de conhecimento em digitalização e curadoria digital e os seus respectivos processos. No corpo de conhecimento em curadoria digital, as áreas de conhecimento tem forte influência dos componentes funcionais do modelo OAIS. A figura 1 apresenta classes da ontologia para objeto digital, pacote de informação e metadados, e a figura 2 apresenta classes para plano de preservação digital.

SUMÁRIO

Figura 3 - Corpos de Conhecimento



Corpo de Conhecimento*	Área de Conhecimento**	Processo de Negócio***
Curadoria Digital	Desenvolvimento de Coleção Digital	Planejar o Acesso e o Uso Planejar o Depósito
	Modelagem de Objeto Digital	Planejar Objeto Digital
	Empacotamento de Objeto Digital	Planejar o Empacotamento do Objeto Digital Gerar Pacote de Submissão Gerar Pacote de Acesso
	Descrição de Objeto Digital	Planejar Metadados Descritivos, Estruturais Técnicos e Administrativos Descrever o Objeto Digital
	Ingestão de Objeto Digital	Gerar Pacote de Submissão Gerar Pacote de Acesso Normalização de Formato Identificação, Caracterização e Validação de Formato de Arquivo
	Armazenamento de Objeto Digital	Armarizar Pacote de Armazenamento Verificar a Integridade dos Pacotes Gerenciar Mídias de Armazenamento Refrescamento de Mídia
	Acesso a Objeto Digital	Recuperação de Mídia Buscar Objeto Digital Entregar Pacote de Acesso
	Preservação Digital	Monitorar a Comunidade e Tecnologia Desenvolver Política de Preservação Digital Planejar a Preservação Digital Normalização de Formato Migração de Formato Migração de Mídia Refrescamento de Mídia
	Infraestrutura Organizacional	Avançar Formatos de Preservação Desenvolver a Missão do Repositório Especificar Processos e Tipologias Documentais Especificar a Estrutura Organizacional

Corpo de Conhecimento*	Área de Conhecimento**	Processo de Negócio***
Digitalização	Seleção	Definir Critérios de Seleção Selecionar Áreas para Digitalização
	Conversão	Definir Diretrizes para Conversão Definir Diretrizes para Escolha de Scanner Planejar a Conversão
	Descrição e Representação	Áreas de Conhecimento e • Modelagem de Objeto Digital • Empacotamento de Objeto Digital • Descrição de Objeto Digital
	Controle de Qualidade	Elaborar Critérios de Avaliação da Qualidade Planejar o Controle de Qualidade
	Produção dos Representantes Digitais	Planejar a Digitalização Receber Material para Digitalização Preparar Material para Digitalização Digitalizar Material Prover Metadados
	Armazenamento de Objeto Digital	Compor Representante Digital Empacotar Representante Digital Controlar a Qualidade
	Acesso a Objeto Digital	Entregar o Pacote do Repres. Digital Desenvolver Material de Digitalização
	Preservação Digital	Elaborar Projeto de Digitalização Executar Projeto de Digitalização
	Infraestrutura Organizacional	Avançar Formatos de Preservação

Áreas e Processos do
Corpo de Conhecimento em
Curadoria Digital

Áreas e Processos do
Corpo de Conhecimento em
Digitalização

SUMÁRIO

CONSIDERAÇÕES FINAIS

Esta seção apresentou abordagens metodológicas para digitalização alinhadas com abordagens tecnológicas para preservação digital, citando documentos técnicos que apoiam essas abordagens. A representação das abordagens metodológicas por meio de corpos de conhecimento apoiam implantadores de empreendimentos de digitalização e preservação digital enfatizando os conhecimentos necessários (áreas de conhecimento) e nas práticas associadas a esses conhecimentos (processos). O uso de ontologia e BFO traz mais clareza na compreensão e na representação de entidades relacionadas à digitalização e à preservação digital, à medida que os recursos envolvidos são descritos por um instrumento rigoroso, formal e semanticamente rico.

REFERÊNCIAS

- Arp.R; Smith, B.; Spear, A. Building Ontologies with Basic Formal Ontology. 2015.<https://mitpress.mit.edu/9780262527811/building-ontologies-with-basic-formal-ontology/>
- CCSDS.Consultative Committee for Space Data System. Audit and Certification of Trustworthy digital Repositories. 2024b <https://public.ccsds.org/Pubs/652x0m2.pdf>
- CCSDS.Consultative Committee for Space Data System. Reference Model for An Open Archival Information System(OAIS). 2024a <https://public.ccsds.org/Pubs/650x0m3.pdf>
- CONARQ. Conselho Nacional de Arquivos. Recomendações para Digitalização de Documentos Arquivísticos Permanentes. 2010 https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/Recomendacoes_digitalizacao_completa.pdf
- Core Trust Seal. CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025 Extended Guidance. 2022 <https://doi.org/10.5281/zenodo.7051096>
- Dappert A.; Enders, M. Digital Preservation Metadata Standards. 2010. <https://www.doi.org/10.3789/isqv22n2.2010.01>

SUMÁRIO

- Dappert, A. DePICT - A conceptual model for digital preservation. 2013 <https://researchportal.port.ac.uk/en/studentTheses/depict>
- Dappert,A.; Farqhhar, A . Modelling Organizational Preservation Goals to Guide Digital Preservation. 2009b <https://doi.org/10.2218/ijdc.v4i2.102>
- Dappert,A.; Farqhhar, A. Significance Is in the Eye of the Stakeholder. 2009a http://www.doi.org/10.1007/978-3-642-04346-8_29
- DILCIS. Digital Information LifeCycle Interoperability Standards Board. E-ARK CSIP: Common Specification for Information Packages. 2021 <https://earkcsip.dilcis.eu/pdf/eark-csip.pdf>
- FIOCRUZ, Manual de Digitalização. 2019 <https://www.arca.fiocruz.br/handle/icict/37187>
- Kennedy, A. et al. Moving Theory into Practice - Digital Imaging Tutorial . 2003 <http://preservationtutorial.library.cornell.edu>
- Kunze, J. *et al.* The BagIt File Packaging Format (V1.0). 2018. <https://www.rfc-editor.org/info/rfc8493>
- Moreau, L. *et al.* PROV-DM: The PROV Data Model - W3C Recommendation. 2013 <https://www.w3.org/TR/2013/REC-prov-dm-20130430/>
- Nascimento *et al.* Guia para digitalização de documentos Versão 2.0. 1006 <https://www.embrapa.br/documents/1355746/1441198/GuiaDigitalizacao.pdf/6394c2e2-be46-4c72-8e81-07247b5f2a38>
- Ören, T. Toward the Body of Knowledge of Modeling and Simulation. 2005. <https://www.site.uottawa.ca/~oren/pubs-pres/2005/pub-0513-MSBOK-IITSEC.pdf>
- Pires, C.; Rocha, R. Finalidades e Atividades da Curadoria Digital na Perspectiva de sua Implantação em uma Instituição. 2020 <https://doi.org/10.36311/1940-1640.2020.v14n4.10857>
- PREMIS Editorial Committee. PREMIS Data Dictionary for Preservation Metadata version 3.0, 2015 <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
- Rocha, 2025. Ontologia Corpo de Conhecimento. 2025a. https://www.ufrgs.br/fabrico/bok/index.php?title=Corpo_de_Conhecimento_em_Digitalização
- Rocha, 2025. Corpo de Conhecimento em Digitalização. 2025b. https://www.ufrgs.br/fabrico/bok/index.php?title=Corpo_de_Conhecimento_em_Curadoria_Digital
- Rocha, 2025. Corpo de Conhecimento em Curadoria Digital, 2025c https://www.ufrgs.br/fabrico/bok/index.php?title=A_Ontologia

12

*Caterina Marta Groposo Pavão
Ana Julia Lopes
Lucia Helena Cunha Vidal*

DADOS QUE FALAM POR SI:

**AValiação da Qualidade dos
DADOS PARA uma Pesquisa
Transparente e Reprodutível**

DOI: 10.31560/pimentacultural/978-85-7221-475-9.12

SUMÁRIO

COMO CITAR:

PAVÃO, Caterina Marta Groposo; LOPES, Ana Julia; VIDAL, Lucia Helena Cunha. Dados que falam por si: avaliação da qualidade dos dados para uma pesquisa transparente e reprodutível. In: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

Dados são uma necessidade absoluta tanto no ambiente empresarial e de negócios como na área científica e de pesquisa. As organizações hoje dependem de dados para cada decisão e consideram os dados como um ativo significativo. Enquanto cientistas de dados e pesquisadores das diversas áreas do conhecimento lutam por dados confiáveis para alimentar suas soluções e estudos científicos, a qualidade dos dados está assumindo uma prioridade cada vez maior. À medida que nos tornamos mais dependentes de dados a importância de monitorar sua qualidade aumenta proporcionalmente. A avaliação de qualidade dos dados (AQD) pode ser implementada para ajudar empresas, instituições e pesquisadores a atingir seus objetivos.

Pesquisas confiáveis exigem dados de qualidade, no entanto, medir a qualidade dos dados é uma etapa crítica em qualquer tarefa de análise de dados. A qualidade dos dados (QD) tem sido um tópico de estudo por muitos anos e continua a receber muita atenção, pois o valor dos dados é altamente dependente de sua qualidade e facilidade de uso (Reda; Benabdellah; Zellou, 2023).

A qualidade dos dados é uma preocupação essencial nas instituições e para os pesquisadores, dados de baixa qualidade

SUMÁRIO

podem custar muito tempo e recursos para serem corrigidos. Esse contexto leva as instituições e os pesquisadores a tomarem medidas e implementar verificações rotineiras de qualidade de dados. Da mesma forma, a confiança nos dados é essencial para a utilização dos dados, dados incompletos, duplicados, redundantes ou imprecisos podem ser resultantes de erros humanos, ferramentas isoladas, múltiplas transferências e estratégia de dados inadequada.

A avaliação da qualidade dos dados tende a se concentrar estritamente na precisão, mas a precisão dos dados não tem sentido a menos que os dados sejam confiáveis, consistentes, completos, atualizados, acessíveis, compreensíveis e relevantes (Stvilia; Lee, 2024). Ademais, os autores acrescentam que existe um “ente” que tem um papel fundamental na avaliação e aferição da qualidade dos dados e da informação, os repositórios digitais, que garantem a autenticidade, a integridade e a privacidade dos dados.

De certa forma, a qualidade dos dados se refere à integridade dos dados que, não só exige que os dados sejam precisos, consistentes e completos, mas também que estejam em contexto. Assim, a integridade dos dados é a garantia da qualidade e consistência dos dados ao longo de seu ciclo de vida completo (SIXSIGMA.US, 2024). Manter a integridade dos dados, para garantir que não ocorrerão mudanças ou alterações não intencionais quando os registros de dados forem modificados, atualizados ou integrados é uma das responsabilidades dos repositórios de dados de pesquisa (RDP). Por sua vez, uma gestão eficaz da QD proporciona ganhos como: dados precisos e confiáveis, análise mais aprimorada, economia de tempo e recursos.

Para que os dados atendam às qualidades mencionadas, é essencial estabelecer e aplicar regras e procedimentos que assegurem sua qualidade. As verificações de validação podem abranger rotinas de limpeza e processos de tratamento de erros, tais como:

SUMÁRIO

validação de dados assegurando integridade no que se refere a valores, intervalos, formatos, entre outros; limpeza de dados (padronização de dados por meio de rotinas de transformação, análise, correspondência e mesclagem); e enriquecimento de dados, melhorando a qualidade ao complementar informações ausentes com dados de outras fontes (SIXSIGMA.US, 2024).

Para o gerenciamento eficaz da QD de um repositório de dados pode ser necessário utilizar ferramentas para automatizar e agilizar processos de avaliação da qualidade dos dados. De acordo com o Guia desenvolvido pela SIXSIGMA.US (2024) podem ser ferramentas para criar perfil de dados, analisar o conteúdo, a estrutura e as métricas dos ativos de dados para descobrir problemas de qualidade nos dados. Ferramentas de limpeza de dados (ou depuração de dados) padronizam formatos de dados, corrigem dados inválidos ou incompletos, removem duplicatas e melhoram a qualidade geral. Essas ferramentas usam técnicas como análise sintática, correspondência, transformação de dados e enriquecimento. As ferramentas de governança de dados estabelecem políticas, processos, funções e métricas para garantir que os dados sejam confiáveis e acessíveis. Essas ferramentas permitem gerenciamento de metadados, controles de segurança e auditoria.

Uma das áreas nas quais a observância da qualidade dos dados é fundamental é a área da Saúde. Os dados de saúde podem apresentar problemas de qualidade, como incompletude, imprecisão e falta de adequação. A avaliação da qualidade dos dados é primordial para permitir o uso adequado e a análise dos estudos, bem como compreender e quantificar as limitações dos estudos clínicos e translacionais. O uso de métodos de avaliação da qualidade de dados baseados em consenso e padronizados pode garantir a aceitação da comunidade, a transparência e a consistência dos resultados, levando a uma melhor compreensão das limitações do estudo, identificação das incertezas e reprodutibilidade dos resultados (Rajan, 2019).

SUMÁRIO

As colocações acima, sobre a importância de garantir a qualidade dos dados, nos mostram também a íntima relação com a reprodutibilidade, o compartilhamento e a reutilização de dados. De acordo com Stvilia *et al.* (2024) os RDP estão na vanguarda do enfrentamento dos problemas a serem solucionados. As práticas de garantia de qualidade de dados nos repositórios podem resultar na concepção de modelos de garantia de qualidade de dados para os RDP.

Mas, é necessário compreender como os repositórios e a comunidade entendem a qualidade dos dados, como garantem a qualidade dos mesmos e quais os meios computacionais disponíveis para executar a tarefa de controle da qualidade dos dados. Sem esquecer e levando em conta que a qualidade dos dados é contextual, diferentes disciplinas e áreas de pesquisa podem usar diferentes metodologias de pesquisa, bem como modelos e fontes de referência distintas para avaliar a qualidade dos dados.

Este estudo sobre AQD baseia-se na literatura recente da área, como os estudos de Stvilia e Lee (2024), Stvilia *et al.* (2024), Stvilia; Lee e Gibradze (2023), Stvilia e Gibradze (2022) e Stvilia (2021), entre outros. Também, aborda as questões estreitamente relacionadas com a AQD como, compartilhamento e reuso dos dados e reprodutibilidade das pesquisas, as dimensões da qualidade dos dados e métodos, técnicas e ferramentas necessárias para avaliar a qualidade dos dados.

DADOS DE PESQUISA

Os dados sempre foram elementos importantes nas pesquisas, especialmente nas pesquisas científicas. São informações muitas vezes “brutas” que precisam ser tratadas para posterior disseminação. O paradigma intensivo de dados da ciência e da pesquisa tem sido explicado como uma tendência emergente que influencia o trabalho de informação.

SUMÁRIO

Com o surgimento das tecnologias digitais, e das possibilidades de busca e recuperação da informação por meio de mecanismos digitais mais aprimorados surgem pesquisas que enfatizam a Ciência de Dados. Esta ciência reúne múltiplos aspectos da informação a partir de seus dados, contando com uma equipe multidisciplinar de profissionais como, por exemplo, cientistas de dados, estatísticos, analistas de dados, programadores, matemáticos e bibliotecários (Reis, 2019).

Dados podem ser definidos como “qualquer” informação que pode ser armazenada em formato digital, incluindo texto, números, imagens, vídeo, filmes, áudio, *software*, algoritmos, equações, animações, modelos, simulações etc. A capacidade de armazenar grandes quantidades de dados por redes de alta largura de banda, suplementadas por outros componentes da infraestrutura de tecnologia da informação e comunicação (TIC) altamente desenvolvida, desencadeou enorme interesse em dados de pesquisa (Boyd; Crawford, 2012). O espectro de atividades relacionadas a dados inclui compartilhamento de dados, gerenciamento de dados, curadoria de dados e citação de dados.

O entendimento sobre o termo “dado” pode variar consideravelmente entre pesquisadores e áreas do conhecimento. Semeler e Pinto (2019) argumentam que a palavra “dado” significa uma única peça de informação. Para os autores podem ser fatos ou estatísticas reunidas para referência ou análise de algo; também podem ser uma coleção diversa incluindo imagens, fluxo de vídeos ou áudio, *software*, algoritmos etc. Hjørland (2018), acrescenta que os dados possuem seu conceito utilizado em vários campos como banco de dados, mineração de dados, dados vinculados, arquivo de dados, metadados e *Big Data*.

Os dados sempre estiveram presentes nas pesquisas, principalmente de cunho científico. Com a expansão do uso das tecnologias e possibilidades de buscar mecanismos mais elaborados

SUMÁRIO

para realizar pesquisas, e a capacidade de armazenar grandes quantidades de dados em redes de infraestrutura digital, tem despertado o interesse nos dados de pesquisa em pesquisadores de diversas áreas, inclusive das ciências naturais e sociais, das artes e humanidades (Reis, 2019; Koltay, 2017).

Com a consolidação da Ciência Aberta os dados de pesquisa começam a ser reconhecidos como elementos importantes, principalmente a partir do momento em que documentos acadêmicos se tornam objetos digitais complexos, onde novos serviços e ferramentas são desenvolvidos para produzir, publicar, disseminar e gerenciar esses conceitos de mídia acadêmica, dando a entender que os padrões e fluxos da comunicação científica estão rapidamente se ajustando às novas normas de publicação e compartilhamento de dados de pesquisa (Sayão; Sales, 2020; Silveira, 2021).

Conceituando o que são dados de pesquisa, nos apoiamos do relatório da OCDE publicado em 2007, que define os dados de pesquisa como “[...] registros factuais (pontuações numéricas, registros textuais, imagens e sons) usados como fontes primárias para pesquisa científica e que são normalmente aceitos na comunidade científica como necessários para validar os resultados da pesquisa” (Organização para a Cooperação e Desenvolvimento Econômico, 2007, p. 13, 26, tradução nossa).

Sant’Anna, Dias e Maculan (2019) declaram que o acesso aos dados das pesquisas contribui para que ações de organização sejam implementadas, para em seguida os dados serem compartilhados e novo conhecimento produzido. Verifica-se, assim, um novo fenômeno, em que os dados das pesquisas assumem importância considerável. Além desse aspecto, salienta-se a quantidade de dados produzidos pelas instituições de ensino e pesquisa e que possuem potencial de reuso. Ter os dados de pesquisa disponíveis para reuso provoca a necessidade de níveis crescentes de infraestrutura de gestão e de tecnologia, de modo que o reuso dos dados de pesquisa

SUMÁRIO

só se torna possível se forem recuperados, acessados, legíveis e estáveis a longo prazo.

Na comunidade científica, há esforços para a conscientização e motivação dos pesquisadores para compartilhar e reutilizar dados de pesquisa com o intuito de potencializar a reprodutibilidade das pesquisas. Diante disso, a sensibilização de pesquisadores tem sido tema recorrente na comunidade científica internacional e, na maioria das vezes, está relacionada ao contexto da Ciência Aberta (Monteiro; Sant'Ana, 2023).

Araújo, Dias e Autran (2021) em pesquisa sobre o compartilhamento de dados em contexto de Ciência Aberta, observaram que a temática de compartilhamento dos dados está entrelaçada com a sua adequada gestão e publicação, visto que, na ausência de um gerenciamento efetivo e de práticas que tornem os dados públicos, a reutilização deles é afetada, logo, prejudica as novas descobertas científicas.

O compartilhamento de dados é valioso por si só. O compartilhamento de dados, na medida do possível, reduz a necessidade de outros pesquisadores coletar dados semelhantes, permite análises combinadas com outros conjuntos de dados e pode criar recursos importantes para estudos futuros imprevistos. Os conjuntos de dados podem reter seu valor por um período considerável, dependendo da área e do campo de estudo. A visão do compartilhamento de dados como inerentemente valioso não está isenta de desafios. Na verdade, retirar dados de seu contexto original pode ser problemático e levar à reutilização inapropriada (Peng; Hicks, 2021).

Destaca-se que esse fato pode ser superado a partir de iniciativas de boas práticas que se tornem comuns entre os pesquisadores, iniciativas essas, que já estão acontecendo nas universidades americanas e europeias por meio de políticas e atividades de capacitação para suas comunidades acadêmicas.

COMPARTILHAMENTO E REUSO DE DADOS

O sucesso da ciência intensiva em dados está intrinsecamente ligado ao comportamento dos pesquisadores que a praticam. Entre os aspectos fundamentais dessa abordagem científica, destaca-se o compartilhamento de dados, um conceito essencial que depende diretamente da postura e das práticas dos pesquisadores para se consolidar e gerar resultados significativos. De acordo com Koltay (2015) o compartilhamento, neste contexto, é a liberação de dados de pesquisa para uso por outros.

Vários fatores podem motivar pesquisadores a compartilhar seus dados, como por exemplo, permitir o acesso aos dados por outros pesquisadores, e pode ser o pré-requisito para receber financiamento para projetos e pesquisas, conforme estabelecido por diferentes agências de financiamento com um grau variável de rigor, de acordo com o país, instituição ou disciplina.

O compartilhamento de dados facilita sua reutilização por outros pesquisadores, sendo uma prática amplamente incentivada pelos financiadores das pesquisas. Embora nem todos os dados sejam apropriados para compartilhamento, a tendência predominante é adotar a abertura como padrão, impondo restrições apenas quando justificadas por razões legais, éticas ou comerciais específicas (Research Data Oxford, 2024).

Monteiro e Sant'Ana (2024) afirmam que embora o movimento da Ciência Aberta promova a abertura dos dados de pesquisa, alguns pesquisadores têm optado por não compartilhar seus dados por variados motivos, como: competitividade, precedência em descobertas científicas ou prioridade de autoria, o fato de não perceberem benefícios para a sua carreira ao serem citados pelos seus pares, por considerarem que outros pesquisadores terão mais benefícios do que eles (mais publicações), por não pensarem na

SUMÁRIO

responsabilidade ética de disponibilizar os dados ou por não compreenderem as normas institucionais.

Os objetivos da Ciência Aberta incluem maior colaboração científica interdisciplinar, acessibilidade de dados e maior reprodutibilidade e transparência do trabalho científico. Estes dependem de uma maior partilha de dados científicos e de dados em acesso aberto. A partilha de dados é cada vez mais vista como um motor essencial na direção em que a ciência se move a nível mundial e entre disciplinas (Tenopir *et al.*, 2020).

O compartilhamento de dados entre pesquisadores é essencial em diversificadas áreas e contribui para acelerar a colaboração científica, conduzir estudos comparativos, permitir pesquisas em larga escala, verificar e validar resultados de pesquisas, aumentar a eficiência e a qualidade da pesquisa, contribuir em diversos campos de atuação e facilitar a reutilização e a reprodutibilidade de dados por terceiros (Kim; Zhang, 2015).

Monteiro (2021) corrobora que os esforços significativos de diferentes agentes são observados para o encorajamento e conscientização da comunidade científica para compartilhar e reutilizar dados de pesquisa com o intuito de melhorar a reprodutibilidade e avançar nas descobertas científicas, mas com a segurança que os dados não serão copiados, modificados ou distribuídos ilegalmente e que os autores serão reconhecidos e citados.

Práticas sólidas de gerenciamento de dados são necessárias para atingir os objetivos da Ciência Aberta. As melhores práticas em gerenciamento de dados exigem que os cientistas compartilhem seus dados depositando os conjuntos de dados em repositórios confiáveis que podem ser temáticos, governamentais ou institucionais, fornecendo metadados que tornem seus dados localizáveis e citando ou reconhecendo sua preservação e reutilização.

SUMÁRIO

De acordo com Silva, Prado e Araújo (2024) o compartilhamento pode ocorrer de diferentes formas, por exemplo, entre pesquisadores, depósito em repositórios, publicação de artigos de dados etc. Para o reuso desses dados deve-se levar em consideração aspectos técnicos como a forma de estruturação e organização dos dados, e aspectos éticos, como a disponibilidade do consentimento de uso, o anonimato e o reconhecimento do trabalho de terceiros.

Caregnato, Rocha e Gabriel Junior (2021) ainda complementam que a defesa do compartilhamento de dados de pesquisa e a luta pela superação de obstáculos à sua reutilização são motivadores da busca por diretrizes que orientem as ações de gestão necessárias, e os princípios FAIR são um marco importante nesse esforço e na ampliação do valor dos dados para seu reuso.

Em 2014, em um workshop na Universidade de Leiden, os princípios FAIR, acrônimo de *Findable, Accessible, Interoperable e Reusable*, foram criados para orientar o processo de compartilhamento de dados, tornando-os localizáveis, acessíveis, interoperáveis e reutilizáveis por humanos e máquinas (Wilkinson *et al.*, 2016). Esses princípios foram recomendados por periódicos, comunidades científicas e agências de fomento e estão se tornando os princípios orientadores essenciais para estratégias de compartilhamento de dados.

As diretrizes de dados FAIR foram desenvolvidas, também, para facilitar e aprimorar a reutilização de dados de pesquisa. Em suma, os dados compartilhados devem ser localizáveis e acessíveis, o que significa que os dados não podem ser armazenados e/ou ocultos em "gavetas de arquivo". Depois de encontrar os dados e acessá-los, o pesquisador deve ser capaz de reutilizá-los. Fouad *et al.* (2024) declara que isso significa que os dados precisam atender a um conjunto mínimo de padrões e fornecer informações suficientes para torná-los mais compreensíveis por mais pesquisadores do que apenas os criadores dos dados. Para uma reutilização eficaz dos dados, eles devem ser interoperáveis, o que significa que devem ter

SUMÁRIO

uma estrutura esperada que facilite a capacidade das máquinas de entender a estrutura dos dados, ao mesmo tempo em que oferecem suporte à sua reutilização por humanos.

Destaca-se que a gestão de dados de pesquisa tem contribuído e favorecido o reuso de dados, disponibilizando aos pesquisadores dados reutilizáveis, isto é, ao publicarem os dados, os criadores dos dados fornecem uma documentação de metadados semânticos que asseguram a sua autenticidade, reconstrução correta, integridade e precisão (Sayão, Sales, 2015).

Além disso, destaca-se que, a fim de que os dados possam ser reutilizados após o compartilhamento, é essencial que sejam devidamente preservados. A preservação de dados garante que eles permaneçam acessíveis para estudos futuros, análises adicionais e reutilização, permitindo a conservação de informações históricas e culturais ao longo do tempo, por exemplo (Research, 2024). A curadoria ativa transforma dados armazenados para uso imediato em dados preservados com um propósito de longo prazo, assegurando que os pesquisadores possam acessá-los e utilizá-los após o fim do estudo. Dessa forma, a preservação de dados amplia as possibilidades de desenvolvimento de pesquisas futuras, tanto pelos pesquisadores originais quanto por outros pesquisadores.

Quando os dados são compartilhados com as devidas condições de uso e preservados de maneira adequada, eles podem ser reutilizados, o que facilita a reprodutibilidade. Em síntese, o compartilhamento, reutilização e preservação de dados não apenas facilitam a reprodutibilidade das pesquisas, mas também aumentam a transparência e a confiabilidade dos resultados. Para que os dados sejam reutilizáveis, eles precisam ser bem documentados, organizados e preservados, e a reprodutibilidade se beneficia dessa prática ao reutilizar dados de alta qualidade.

SUMÁRIO

Feinberg *et al.* (2020) destacam que, para o trabalho científico ser reprodutível, os processos e resultados da pesquisa devem ser compartilhados. Logo, o compartilhamento de dados está interligado ao conceito de reprodutibilidade, e a reutilização de dados é promovida como um objetivo principal do compartilhamento de dados. O compartilhamento de dados foi identificado como um componente fundamental da colaboração científica. Os dados que foram preservados e disponibilizados para acesso público facilitam a reprodutibilidade ao permitir estudos de reanálise e replicação e permitem que novas combinações de dados surjam. A reutilização de dados pode ter impactos econômicos, pois reduz o custo da coleta de dados redundantes ao disponibilizar os dados coletados a outros pesquisadores, proporcionando oportunidades de colaboração e coautoria.

Em sua pesquisa, Goble, Roure e Bechhofer (2013) identificaram que dada a tensão entre o trabalho envolvido no compartilhamento de dados e seus incentivos, os pesquisadores se voltaram para as principais razões de compartilhar dados e como encorajar ou impor práticas reproduzíveis. Os autores identificaram que as principais práticas de compartilhamento de dados, como métodos de comunicação e descrição de metadados, tornam-se essenciais para facilitar a reprodutibilidade.

Tornar a ciência reproduzível um padrão envolve, o que Stodden e Miguez (2013) chamaram de “mudança cultural” e requer uma mudança nas políticas de instituições científicas centrais, dos editores de periódicos e agências de financiamento.

REPRODUTIBILIDADE DAS PESQUISAS

A Reprodutibilidade se refere à capacidade de reproduzir ou recriar os resultados de um experimento, análise ou estudo utilizando os mesmos métodos, dados e procedimentos que foram usados

SUMÁRIO

originalmente. Em outras palavras, se outro indivíduo repetisse o mesmo experimento ou análise, deveria ser capaz de obter resultados semelhantes aos do pesquisador original (Salomão, 2023).

A mesma autora ainda explica que a reprodutibilidade é um aspecto essencial do método científico uma vez que ajuda a estabelecer a validade e a confiabilidade dos resultados da pesquisa, permitindo que outros pesquisadores validem os resultados e conclusões, desenvolvam e verifiquem ou refutem os resultados originais. Portanto, para garantir a reprodutibilidade, os pesquisadores precisam tornar seus dados, códigos e métodos abertamente disponíveis e bem documentados, e garantir que seus experimentos e análises possam ser replicados por outros.

A reprodutibilidade refere-se à capacidade de um pesquisador de duplicar os resultados de um estudo anterior utilizando o mesmo material como os usados pelo pesquisador original. Isto é, um segundo pesquisador pode usar os mesmos dados brutos para construir a mesma análise e implementar a mesma análise estatística na tentativa de produzir os mesmos resultados. A reprodutibilidade é um mínimo necessário para que uma descoberta seja crível e informativa (Goodman; Fanelli; Ioannidis, 2016).

Em seu estudo sobre as definições de “pesquisa reprodutível”, Barba (2018) destaca que algumas áreas do conhecimento tratam os termos “reprodutível” e “replicável” como sinônimos, enquanto outras disciplinas os empregam com significados opostos.

A respeito disso, Peng e Hicks (2021) declaram que existe uma diferença fundamental entre reprodutibilidade e a replicação. A reprodutibilidade refere-se à capacidade de outros pesquisadores obterem os mesmos resultados de um estudo quando usarem os mesmos dados e seguirem os mesmos métodos, procedimentos e código originais. Nesse caso, as diferenças entre os resultados de pesquisadores distintos devem ser mínimas, exceto por alguma tolerância para diferenças na precisão da máquina.

SUMÁRIO

Já a replicação, por sua vez, implica na capacidade de um estudo ser repetido de forma independente, sob condições similares, produzindo resultados consistentes com os originais. Dois pesquisadores conduzindo exatamente o mesmo experimento devem, em teoria, diferir apenas por uma quantidade quantificada e pelo desvio padrão dos dados (Peng; Hicks, 2021).

A replicação, ou a repetição completa de um estudo para verificar seus resultados, é a forma mais eficaz de confirmar descobertas e minimizar erros, omissões e fraudes. No entanto, devido aos custos elevados de reproduzir pesquisas ou reanalisar fontes de dados, a replicação nem sempre é viável. Como alternativa, propõe-se a reprodutibilidade como um padrão mínimo para aumentar a confiabilidade dos resultados reportados. Uma pesquisa é considerada reprodutível quando seus dados são acessíveis e os métodos de gerenciamento e análise são descritos de forma clara e completa (Harris *et al.*, 2018).

Outrossim, a adoção de práticas de pesquisa reprodutíveis acelera a descoberta científica, promove maior troca de ideias entre cientistas e reduz o desperdício de pesquisa, de forma que os artigos de pesquisa com dados compartilhados têm menos erros, aqueles com dados compartilhados ou código compartilhado são mais citados e as fontes de dados compartilhadas são a base para mais publicações (McKiernan *et al.*, 2016).

Com o movimento da Ciência Aberta, a pesquisa reprodutível tornou-se objetivo generalizado em comunidades de pesquisa, círculos políticos e órgãos de financiamento de pesquisas. O entendimento é que práticas de pesquisa abertas e reprodutíveis permitem a reutilização científica, acelerando projetos e descobertas futuras em qualquer disciplina. Com o intuito de tomar medidas concretas em busca desses objetivos, tem havido discussões acompanhadas para tornar os resultados de pesquisa científica abertos e reprodutíveis (Chen *et al.*, 2019).

SUMÁRIO

Samuel e König-Ries (2021) afirmam que a reprodutibilidade sempre foi importante para a comunidade científica, pois apoia a extensão e a construção em cima dos trabalhos de outros pesquisadores promovendo o progresso científico, e possibilitando aos pesquisadores conduzirem melhores estudos, permitindo que eles verifiquem seus próprios resultados e de outros pesquisadores, aumentando assim a confiança no estudo científico.

Embora a reprodutibilidade tenha sido um princípio importante da boa ciência por muito tempo, a ênfase recente em realizar a reprodutibilidade na prática é nova. O recente ímpeto em direção à reprodutibilidade, portanto, surge como uma ampla mudança cultural, com o objetivo de alterar as estruturas institucionais e de incentivo existentes e mudar a maneira como a ciência é feita no nível da prática. Contudo, a reprodutibilidade permanece como um dos desafios centrais na ciência contemporânea (Feinberg *et al.*, 2020).

A "crise da reprodutibilidade" ganhou destaque no cenário científico após um estudo publicado pela revista Nature em 2016 (Baker, 2016), que entrevistou 1.576 pesquisadores. Os resultados revelaram que aproximadamente 90% dos cientistas reconhecem a existência dessa crise caracterizada pela dificuldade em reproduzir resultados de estudos publicados, causando impactos no progresso científico e na confiabilidade nos resultados científicos.

A falta ou a impossibilidade de reprodutibilidade das pesquisas científicas refere-se a situações em que tentativas de reproduzir as descobertas de um estudo falham em produzir resultados consistentes e muitos fatores podem contribuir para esse déficit de reprodutibilidade na pesquisa científica.

De acordo com Salomão (2023) alguns fatores que cooperam para a carência de consistência na reprodução das pesquisas científicas são a disponibilização de relatórios, muitas vezes incompletos ou imprecisos, que dificultam para outros pesquisadores reproduzir os resultados; estudos conduzidos com tamanhos de amostra

SUMÁRIO

pequenos, podendo prejudicar que outros pesquisadores repliquem as descobertas; ou ainda, intercorrência metodológica, quando há diferenças nos métodos ou materiais específicos utilizados na pesquisa que podem afetar o resultado. Isso impossibilita determinar se uma falha na replicação é devido a uma falta genuína de reprodutibilidade ou simplesmente de diferenças metodológicas.

Por outro lado, demonstrar reprodutibilidade mostra integridade, inspira confiança e incentiva a reutilização. Isso significa que o pesquisador receberá mais atenção para cada artigo individual, exercerá um impacto maior em sua comunidade de pesquisa e contribuirá para um ecossistema de pesquisa mais eficiente.

Morton (2022) publicou algumas práticas de Ciência Aberta que melhoram a reprodutibilidade e apoiam a confiança na ciência. Entre as práticas citadas pela autora estão:

- a. compartilhar métodos abertos - é difícil reproduzir resultados (muito menos adaptar uma metodologia para reutilização) com base apenas nas informações de um artigo de pesquisa. Não importa se os métodos incluem protocolos, código ou algo mais, torná-los acessíveis inspira confiança, facilita a reutilização e estende a vida útil do trabalho;
- b. documentar e relatar completamente os materiais - os materiais são tão importantes para a reprodutibilidade quanto os procedimentos, protocolos e ferramentas analíticas usadas na condução de um estudo;
- c. publicar dados abertos em um repositório público - dados abertos fornecem os detalhes necessários para que pesquisadores validem, repliquem e reproduzam os resultados uns dos outros. Depositar um conjunto de dados compatível com os princípios FAIR em um repositório garante que os dados permaneçam acessíveis como uma parte permanente do registro científico.

SUMÁRIO

Vale salientar que a publicação dos dados não promove apenas a reprodutibilidade das pesquisas, mas possibilita, também, o seu reuso, onde o valor atribuído ao estudo está relacionado ao potencial dos dados serem reinterpretados em outras áreas e contextos diferentes daqueles que originalmente os gerou, estabelecendo, assim, novos padrões de trabalho cooperativo independente de barreiras geográficas ou disciplinares (Piwowar; Vision, 2013; Ferreira, 2024).

Dados de pesquisa, definidos como informações coletadas ou geradas, usadas como evidência para descobertas de pesquisa originais, tornaram-se um componente vital do registro acadêmico e um ativo primário de novas pesquisas. O aumento do valor dos dados de pesquisa e as preocupações com uma “crise de reprodutibilidade” levaram financiadores e periódicos a exigirem o compartilhamento de dados como condição de financiamento de bolsas e publicação.

Com isso, os repositórios de dados são considerados um local primário para o compartilhamento de dados, uma vez que implementam a administração sistemática para promover a curadoria, disseminação, acesso e preservação de dados de pesquisa. No entanto, os dados publicados só serão reutilizados se um pesquisador confiar em sua qualidade (Pimentel *et al.*, 2019; Stodden; Seiler; Ma, 2018; Assante *et al.*, 2016).

Os problemas de qualidade de dados podem prejudicar significativamente a reprodutibilidade da pesquisa, o compartilhamento de dados e a reutilização. Na vanguarda do tratamento de problemas de qualidade de dados estão os repositórios de dados de pesquisa (RDP), essenciais para garantir a qualidade dos dados, seguindo padrões como os princípios FAIR e normas internacionais como a ISO 8000 e ISO 9000 (Stvilia *et al.*, 2025).

Definir e avaliar a qualidade dos dados pode ser uma tarefa complexa, pois varia dependendo da finalidade específica para a qual os dados são destinados e uma abordagem padronizada para

SUMÁRIO

avaliação da qualidade dos dados pode não ser adequada em todos os casos, pois diferentes usos dos dados podem ter critérios e considerações de qualidade distintos. Melhorar e sinalizar a qualidade dos dados de pesquisa aprimorando a curadoria de dados, a integridade do código e os incentivos de publicação de dados em repositórios de dados contribuirá para a transparência, reprodutibilidade e reutilização de produtos de pesquisa (Trisovic *et al.*, 2021).

QUALIDADE DOS DADOS

A qualidade dos dados (QD) é um dos pilares fundamentais Ciência de Dados e, portanto, também nas pesquisas científicas, uma vez que as análises e modelos gerados dependem diretamente da confiabilidade e precisão das informações utilizadas. De acordo com Strong, Lee e Wang (1997), a QD pode ser definida como a capacidade de um conjunto de dados de atender aos requisitos dos usuários no contexto de uma aplicação específica. Em outras palavras, dados de alta qualidade são aqueles que garantem confiança nos resultados das análises e promovem a consistência em processos baseados em informações.

A Ciência de Dados por ser uma área interdisciplinar envolve a utilização de métodos, algoritmos, processos e sistemas para extrair conhecimento e *insights* de dados em diversas formas, seja estruturado ou não estruturado e para utilização em diversas áreas do conhecimento. Com a crescente quantidade de dados gerados no mundo digital, tornou-se fundamental entender e analisar esses dados, com o objetivo de apoiar decisões, identificar padrões e tendências e fazer previsões com base em dados históricos (Davenport; Patil, 2012; Provost; Fawcett, 2013).

SUMÁRIO

A importância da QD está ligada ao fato de que algoritmos e técnicas analíticas são sensíveis a erros nos dados. Dados de baixa qualidade podem introduzir viés, afetar a generalização dos modelos e resultando em decisões erradas (Pagano *et al.*, 2023). Isso ocorre porque dados imprecisos, incompletos ou distorcidos fornecem informações inadequadas, que influenciam diretamente a capacidade do modelo de aprender padrões reais e relevantes. Muitos modelos de avaliação da qualidade dos dados (AQD) têm sido estudados e implementados nos últimos anos.

Com a crescente implementação de *Big Data* e *Machine Learning*, surgiram modelos como o de Olson (2003) que reforçam a importância da validação e do monitoramento contínuo da qualidade dos dados. O autor argumenta que, para garantir dados de alta qualidade, é essencial estabelecer processos robustos de controle que envolvem a validação de entradas, como também a verificação da integridade e a manutenção contínua da qualidade (Olson, 2003). Tais processos são aplicados desde a entrada de dados nos sistemas até a sua utilização em modelos analíticos, evitando que dados errados ou imprecisos comprometam as conclusões e decisões baseadas neles.

Outra abordagem relevante, no contexto mais atual, é a desenvolvida por Abedjan, Golab e Naumann (2016), que analisam a QD especificamente em ambientes de *big data*. Os autores destacam que, devido ao grande volume, diversidade e velocidade com que os dados são gerados, modelos de qualidade dos dados precisam ser dinâmicos e escaláveis, além de serem capazes de lidar com dados em tempo real.

O modelo proposto por Abedjan, Golab e Naumann (2016) sugere a implementação de ferramentas de monitoramento e limpeza automática, essenciais para lidar com as incertezas dos dados não estruturados ou dispersos em sistemas complexos de *big data*, como os encontrados em *Data Lakes* e ambientes de nuvem.

SUMÁRIO

Destaca-se ainda, que a governança de dados tem ganhado cada vez mais importância como parte da estrutura de modelos de qualidade. A governança de dados envolve a definição de políticas e responsabilidades dentro de uma organização, com o objetivo de garantir que os dados sejam tratados de forma ética, segura e eficaz. Segundo um estudo publicado por SixSigma.us (2023), a governança bem estruturada é fundamental para assegurar que as práticas de qualidade sejam seguidas e que as decisões baseadas em dados sejam as mais precisas e confiáveis possíveis.

O viés introduzido por dados de baixa qualidade pode ser particularmente problemático, pois ele pode refletir preconceitos ou desigualdades nos dados originais, como no caso de dados que não representem adequadamente todas as populações. Isso pode levar a resultados discriminatórios e injustos, afetando negativamente a tomada de decisão em contextos críticos como crédito, saúde e justiça (Pagano *et al.*, 2023). Assim, um esforço significativo deve ser empregado na etapa de preparação dos dados, que muitas vezes consome até 80% do tempo de um projeto (Press, 2016).

A implementação de modelos de QD é complexa e essencial. A QD não é apenas uma questão técnica, mas também estratégica, pois impacta diretamente a eficácia das análises e a confiabilidade das decisões tomadas durante a execução de um projeto com base nas informações analisadas. Instituições e pesquisadores que adotam boas práticas e aplicam modelos eficazes de qualidade podem extrair o máximo valor de seus dados e se destacar no competitivo cenário atual, seja no âmbito empresarial ou de pesquisas.

Detectar e corrigir inconsistências em grandes volumes de dados muitas vezes requer processos manuais para interpretar problemas complexos (Kandel *et al.*, 2011). A limpeza de dados, que envolve a correção ou remoção de dados incorretos, incompletos ou duplicados, assegurando que os dados estejam prontos para uso, (Rahm; Do, 2000); validação de entradas, garantindo que as

SUMÁRIO

informações inseridas estejam de acordo com os padrões definidos e com formatos abertos (Wand; Wang, 1996); e monitoramento contínuo para detectar falhas e melhorar os processos de dados ao longo do tempo, prevenindo problemas (Sadiq; Indulska, 2017), são estratégias essenciais para garantir qualidade.

Também, *frameworks* de governança de dados são relevantes ao estabelecer políticas e diretrizes que garantem a gestão eficaz dos dados. Essas práticas incluem os planos de gestão de dados (PGDs), que documentam a origem, rastreiam as alterações e permitem auditorias regulares, elementos essenciais para manter a confiança nos dados ao longo do tempo.

A qualidade dos dados influencia diretamente na confiabilidade e efetividade das análises baseadas em dados. Compreender suas dimensões, impactos e estratégias de garantia permitem extrair valor significativo dos dados. Dados de alta qualidade são indispensáveis para qualquer pesquisa que dependa de informações confiáveis e enfatizar a importância de integrar e manter dados consistentes, completos e atualizados torna-se imprescindível. A complexidade crescente dos dados, incluindo sua origem e as diferentes tecnologias utilizadas para armazená-los e analisá-los, coloca desafios na garantia de sua qualidade.

Por fim, uma definição curta e pragmática de qualidade de dados é a fórmula *adequado para o propósito*, ou seja, uma orientação para o propósito ou intenção de uso. Essa ideia central origina-se da garantia de qualidade em processos de produção industrial, mas a fórmula, *adequado para o propósito* ou *adequado para uso* também é usada em muitos campos da ciência. Dessa forma, a qualidade dos dados é definida de forma abrangente, mais aberta, como a totalidade de propriedades e características dos dados que os tornam adequados para um propósito específico (German, 2020).

DIMENSÕES DE QUALIDADE DOS DADOS

Os estudos sobre qualidade dos dados remontam a quase 30 anos atrás, quando Wang e Strong (1996) publicaram uma estrutura amplamente utilizada para avaliar a qualidade dos dados. A estrutura pensada pelos autores descreve 20 dimensões de qualidade agrupadas em quatro categorias. Era direcionada na perspectiva do usuário e, embora o contexto original fosse para aplicação em dados empresariais, tem sido aplicada no contexto de dados de pesquisa.

Os conceitos de AQD podem ser abordados com foco em diferentes dimensões, práticas e modelos utilizados em repositórios de dados de pesquisa (RDP). A avaliação da qualidade dos dados é um componente importante de cada operação de análise de dados. É o processo de determinar se os dados atendem aos requisitos de informação de um usuário em um caso de uso específico como exposto no estudo de Silvola *et al.* (2016). Envolve medir as dimensões ou critérios de qualidade que são relevantes e comparar as descobertas com as necessidades dos pesquisadores.

Um dos desafios na AQD, de acordo com Silvola *et al.* (2016), surge da natureza contextual da qualidade dos dados. É essencial considerar o contexto específico no qual os dados são usados, pois a AQD é altamente dependente da finalidade pretendida e do cenário específico. Consequentemente, a mesma dimensão de qualidade pode ter relevância diferente e exigir métodos de avaliação distintos em contextos diferentes.

Para Wang e Strong (1996) as dimensões de qualidade de dados são uma coleção de atributos de qualidade que refletem cada uma uma característica ou construção distinta de qualidade de dados, ou seja, cada dimensão diz respeito a um aspecto específico do dado. Corroborando com os autores, as dimensões da QD “[...] são virtudes ou características dos dados usados para definir e/ou comunicar o conceito de qualidade” (Stvilia; Lee, 2024, p. 800, tradução nossa).

SUMÁRIO

As definições de QD são frequentemente complementadas por dimensões que descrevem aspectos gerais da qualidade de dados e critérios que especificam quais características tornam os dados adequados para uso em um determinado contexto (Kindling; Strecker, 2022). Entre as dimensões de qualidade mais comuns, cuja aplicação depende do contexto analítico, Wang e Strong (1996) destacam a precisão, a completude e a acessibilidade.

Ainda, conforme Stvilia; Lee e Gibradze (2023), Stvilia *et al.* (2024) e Stvilia e Lee (2024), as dimensões da QD podem ser classificadas como: a) *dimensões intrínsecas* que se referem às características internas dos dados, como precisão, completude, consistência, atualidade, confiabilidade, relevância e simplicidade, essas dimensões são essenciais para avaliar a QD em si, independentemente do contexto de uso e b) *dimensões no nível do produto* que se referem às características dos dados como um produto, incluindo acessibilidade, usabilidade, conformidade ética e legal, estabilidade, integridade, reputação e autenticidade, essas dimensões são importantes para garantir que os dados sejam úteis e reutilizáveis.

As dimensões intrínsecas da QD e sua definição pode variar de acordo com o ambiente do repositório e os dados depositados. Nos estudos de Stvilia; Lee e Gibradze (2023), Stvilia *et al.* (2024) e Stvilia e Lee (2024) cada dimensão intrínseca é definida da seguinte maneira: a) *precisão*, grau em que os dados representam corretamente o objeto, relação, processo ou evento; b) *completude*, extensão em que os dados estão completos em relação ao que se propõem a representar; c) *consistência*, uniformidade na representação dos dados ao longo do tempo ou entre diferentes fontes; d) *atualidade*, idade dos dados e sua relevância temporal; e) *confiabilidade*, capacidade de verificar a correção dos dados; f) *relevância*, aplicabilidade dos dados para uma determinada pesquisa e finalmente g) *simplicidade*, complexidade cognitiva dos dados.

SUMÁRIO

Em outra perspectiva, para o German Council for Scientific Information Infrastructures (2020), a qualidade dos dados é orientada à demanda, ou seja, da perspectiva do uso dos dados ou consumo dos dados e pode ser diferenciada em qualidade: a) *intrínseca* dos dados: os dados têm uma qualidade própria, por exemplo, por serem, confiáveis e objetivos; b) *contextual*, quando a qualidade dos dados resulta da sua adequação a uma finalidade específica do contexto, mas também, da sua relevância, atualidade e valor agregado; c) *representacional*, se refere a qualidade dos dados quando os dados são concisos, consistentes em seus formatos de apresentação, interpretáveis e fáceis de entender e d) *relacionada ao acesso*, os dados ganham qualidade se forem acessíveis e editável e o acesso a ele é seguro e permanecerá assim no futuro.

Strong, Lee e Wang (1997) acreditam que as dimensões de QD devem incluir além da precisão, características como: objetividade, credibilidade e reputação. Reda, Benabdellah e Zellou (2023) sintetizam, a partir de uma análise sistemática de literatura, as dimensões da avaliação da qualidade dos dados. Além das dimensões precisão, completude e consistência, apontadas por Stvilia; Lee e Gibradze (2023), Stvilia *et al.* (2024) e Stvilia e Lee (2024), acrescentam a confiencialidade das fontes, a recuperabilidade garantida pelos *backups* periódicos, a rastreabilidade de acesso e a compreensibilidade dos dados. Porém, no que se refere à completude dos dados fazem um acréscimo importante, subdividindo-a em: completude de atributos quando todos os dados avaliados devem atender a determinados critérios e completude do registro que mede número de registros que cumprem todos os critérios.

Analisando as dimensões propostas por Zhang, Indulska e Sadiq (2019) verifica-se um desdobramento de três das quatro dimensões apontadas pelos autores para AQD, completude, precisão, validade e consistência, como forma de especificar aspectos capazes de avaliar com maior precisão a qualidade dos dados. A *completude* está relacionada à completude dos atributos obrigatórios (para uma

SUMÁRIO

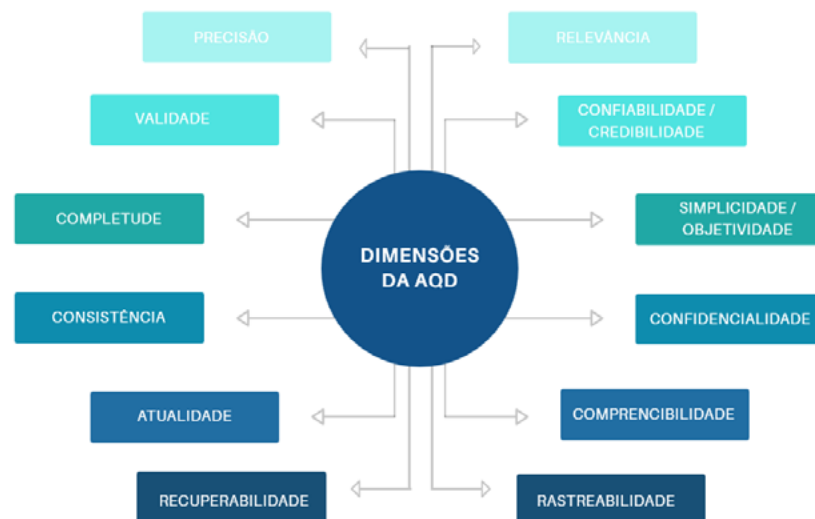
representação completa de uma entidade do mundo real os atributos devem conter valores e não pode ser nulo) e completude dos atributos opcionais (não devem conter valores nulos inválidos). Em relação à *precisão* destacam que os valores dos atributos devem ser precisos de acordo com a linguística e a granularidade. A *conformidade* se refere à conformidade dos dados com as regras de negócios e à conformidade dos metadados.

A dimensão consistência apresenta mais desdobramentos no estudo de Zhang, Indulska e Sadiq (2019), onde são apontados os seguintes aspectos a serem avaliados: *unicidade* (os dados devem ser exclusivamente identificáveis), *não redundância* (os dados devem ser registrados em um único lugar), *consistência semântica* e *consistência de valor* (os valores não devem fornecer instâncias conflitantes ou heterogêneas), *consistência de formato* (quando os formatos de dados são usados consistentemente) e *integridade referencial* (quando os relacionamentos de dados são representados por meio de regras de integridade referencial).

As dimensões da qualidade dos dados elencadas pelos diversos autores citados acima foram resumidas na Figura 1, onde são apresentadas as dimensões intrínsecas dos dados e as características sob a perspectiva do uso dos dados ou consumo dos dados.

SUMÁRIO

Figura 1 – Dimensões da AQD em repositórios de dados de pesquisa



Fonte: elaboração das autoras.

No entanto, quando se refere às características dos dados como um produto, em repositórios de dados de pesquisa o estudo de Stvila *et al.* (2024) aponta que as dimensões acessibilidade, usabilidade e reputação são as mais citadas pelos pesquisadores, seguidas das dimensões de integridade e autenticidade. Ainda, foram evidenciadas dimensões que dizem respeito a componentes da qualidade do sistema como, por exemplo, conformidade ética e legal, acesso, usabilidade, acionabilidade e estabilidade.

De acordo com Reda, Benabdellah e Zellou (2023), a literatura sobre qualidade de dados apresenta definições das dimensões da qualidade de dados, no entanto não há acordo sobre quais dimensões caracterizam a qualidade dos dados ou o que exatamente cada dimensão significa, porque ela é contextual. Essa afirmação vai ao encontro da visão de Bertossi e Rizzolo (2016, p. 6, tradução nossa) onde, “[...] uma avaliação completa da qualidade dos dados não pode ser conduzida isoladamente; [...]. A qualidade depende, entre outras coisas,

SUMÁRIO

do contexto que nos permite dar sentido aos dados e avaliá-los”. A qualidade dos dados deve ser considerada intimamente relacionada ao uso pretendido dos dados e só pode ser avaliada nesse contexto.

Portanto, deve-se considerar classificar os domínios de pesquisa no campo da QD. Na revisão sistemática realizada por Reda, Benabdellah e Zellou (2023), observou-se que vários estudos se concentraram no mesmo tópico, mas adotaram diferentes abordagens e perspectivas, destacando a amplitude e a diversidade da AQD em vários domínios de pesquisa. Alguns dos domínios citados pelos autores no estudo podem ser verificados no Quadro 1, assim como uma descrição desse domínio de pesquisa e as dimensões utilizadas para avaliar a qualidade dos dados.

Quadro 1 – Domínios, descrição e dimensões da AQD

Domínio	Descrição	Dimensões de QD
<i>Big data</i>	Este domínio explora os desafios e técnicas para garantir a qualidade dos dados em conjuntos de dados de grande escala.	Precisão, integridade, consistência, atualidade, completude, distinção, fidelidade, pontualidade, volume, variedade, velocidade e veracidade
Inteligência artificial (IA)	A pesquisa nesta área se concentra no impacto da qualidade dos dados no desempenho da IA e explora estratégias para manter dados de alta qualidade para aplicações de IA.	Precisão, consistência, integridade, atualidade, normatividade, plausibilidade, conformidade e completude
Assistência médica	ADQ em ambientes de saúde ganhou atenção significativa devido ao papel crítico de dados precisos e confiáveis no atendimento ao paciente, tomada de decisão clínica e resultados de saúde.	Precisão, consistência, integridade, atualidade, normatividade, plausibilidade, conformidade e completude.
Internet das coisas (IoT)	É uma área de estudo que fornece perspectivas ilimitadas e é vital para pesquisadores em todo o mundo. É uma revolução tecnológica que alterará a maneira como as pessoas trabalham, pensam, vivem e depende fortemente da precisão dos dados gerados por dispositivos IoT.	Imprecisão, integridade, inconsistência, ambiguidade, incerteza, atualidade, credibilidade, acessibilidade, acesso, segurança, interpretabilidade, precisão, correção, exclusividade, validação, acessibilidade, facilidade de manipulação e representação.
<i>Web</i>	Pela vasta quantidade de dados que agora está disponível na <i>web</i> , ela se tornou a principal fonte de informação. No entanto, mesmo com uma grande quantidade de dados, sua qualidade permanece duvidosa. Assim, avaliar e melhorar o nível de qualidade dos dados fornecidos é cada vez mais importante.	Formato, completude, disponibilidade, recuperabilidade, <i>design</i> e dados conectados.

SUMÁRIO

Domínio	Descrição	Dimensões de QD
Indústria	Existem muitos casos em que a QD pode degradar ao longo da vida de um sistema de produção. Essas infraestruturas de coleta de dados podem ter um desempenho ruim, especificamente em termos de parâmetros de rede, como perdas, atrasos ou carga de tráfego, resultando em uma diminuição na QD.	
Sistemas de informação	Avaliar a QD dentro de um sistema de informação complexo pode ser desafiador devido ao fato de que várias fontes de dados, tanto <i>hardware</i> quanto <i>software</i> , estão envolvidas. Esses desafios incluem analisar e processar uma grande coleção de dados para garantir a QD.	Completude, precisão, acessibilidade, consistência, atualidade e relevância.
Mídias sociais	As mídias sociais surgiram como uma nova fonte de informações úteis, se processados corretamente, podem ajudar a obter <i>insights</i> na tomada de decisões e promover pesquisas em diversas áreas, de modo que avaliar a QD sociais é uma tarefa que depende do contexto.	Confiabilidade, usabilidade, disponibilidade, relevância e qualidade de apresentação.
Contexto geral	Muitos estudos que se concentram na questão da qualidade dos dados em um contexto geral, sem necessariamente serem específicos do domínio.	Integridade, consistência, precisão, singularidade, validade, integridade, atualidade, utilidade, complexidade, validade interna, validade externa, pontualidade e interpretabilidade.

Fonte: adaptado de Reda, Benabdellah e Zellou (2023).

Garantir e melhorar a qualidade dos dados é um desafio permanente em todo o processo de pesquisa e que merece reflexão. O elemento central da qualidade dos dados é a documentação precisa e a divulgação das medidas, ferramentas, *software* utilizados e as etapas processuais para gerar, processar e disponibilizar os dados. A Ciência da Informação e as infraestruturas podem definir princípios técnicos ou padrões como diretrizes para a qualidade. A qualidade científica dos dados é caracterizada por uma pluralidade de critérios. Esses critérios só podem ser definidos e descritos em detalhes pelas próprias comunidades científicas e devem ser integrados à prática de pesquisa. (General, 2020).

MÉTODOS E TÉCNICAS PARA AVALIAÇÃO DA QUALIDADE DOS DADOS

A garantia de qualidade de um conjunto de dados pode ser conceituada como um sistema de atividades que avaliam, intervêm e comunicam a qualidade dos dados e metadados. A atividade de AQD é o processo de identificar problemas de qualidade nos dados e metadados. Envolve a definição de métricas e a aplicação de ferramentas para medir a qualidade dos dados (Stvilia, 2021; Stvilia *et al.* 2007).

Portanto, a garantia da QD demanda uma abordagem sistemática, integrando curadoria ativa, automação, padrões internacionais e modelos de ciclo de vida. Essas estratégias ajudam a manter a integridade, acessibilidade e reutilização dos dados pela comunidade científica e requerem estratégias estruturadas.

A percepção do que constitui dados úteis e de qualidade e/ou quando os dados se tornam úteis pode variar dentro do mesmo processo, disciplina e entre diferentes processos dentro das disciplinas (Stvilia; Gibradze, 2022). Portanto, antes de elaborar métricas, verificações e processos de avaliação da qualidade dos dados e metadados, os repositórios de dados de pesquisa e suas comunidades precisam determinar e chegar a um consenso sobre o que é qualidade dos dados e o que entendem por “adequação para uso” ou “adequação para reutilização” (Stvilia; Lee; Gibradze, 2023). Essa prática de avaliação pode envolver a verificação de valores ausentes ou inválidos, uso de modelos de qualidade de dados específicos da comunidade e avaliação da reputação dos provedores de dados para prever a qualidade dos conjuntos de dados.

Quando os dados não atendem a qualidade esperada pelos usuários ou pela comunidade podem ser necessárias intervenções que são as ações tomadas para corrigir ou melhorar a qualidade dos

SUMÁRIO

dados após a avaliação. Porém, qualquer intervenção deve ser precedida por AQD e metadados. Isso pode incluir a padronização de esquemas de dados, a adição de metadados ausentes ou a remoção de dados redundantes (Stvilia; Lee, 2024). A abordagem de intervenção de qualidade de dados normalmente envolve o pesquisador que deve realizar as alterações necessárias.

Alguns exemplos de intervenções são: correção de erros identificados por meio de análises estatísticas, adição de documentação e metadados para melhorar a compreensão dos dados, rejeição ou remoção de conjuntos de dados que não atendem aos padrões de qualidade.

Ainda, os repositórios de dados podem implementar métodos de comunicação como uma forma de compartilhar informações sobre a qualidade dos dados com os usuários e outras partes interessadas. Isso pode incluir, de acordo com Stvilia *et al.* (2024) relatórios de qualidade, guias de usuário e *feedback* dos usuários. Alguns exemplos de práticas de comunicação são: elaborar documentação sobre os problemas mais frequentes de qualidade fornecendo aos depositantes modelos e guias para abordar lacunas de informação e aprimorar a qualidade dos metadados, preparando conjuntos de dados para publicação, converter conjuntos de dados para formatos de arquivo mais acessíveis ou abertos para aumentar sua utilidade, usar um sistemas de *feedback* para coletar opiniões dos usuários sobre a qualidade dos dados e processos de comunicação proativa de problemas de qualidade para usuários que já acessaram os dados.

Uma das técnicas exploradas para garantir a qualidade dos dados é o envolvimento ativo dos produtores de dados no processo de curadoria (Witt, 2009). Essa prática assegura que os dados sejam compreensíveis e reutilizáveis por outros pesquisadores. O desenvolvimento de perfis de curadoria ajuda a estabelecer políticas claras sobre como os dados devem ser estruturados, documentados e preservados, contribuindo significativamente para a consistência

SUMÁRIO

e confiabilidade dos repositórios. A curadoria proativa é um exemplo eficaz, envolvendo revisões detalhadas antes da publicação dos dados (Hemphill *et al.* 2024). Adicionalmente, automação de avaliações (Rajan *et al.*, 2019) emerge como uma solução que permite a identificação de inconsistências e erros em grande escala.

A adoção de modelos de ciclo de vida, como o *Curation Lifecycle Model* (CLM), oferece um *framework* estruturado para gerenciar dados desde sua criação até a preservação de longo prazo (Sanchez; Vidotti; Vechiato, 2017). Complementarmente, a utilização de padrões de metadados consolidados, como *Dublin Core*, *DataCite* e ISO 19115, facilita a interoperabilidade entre sistemas e melhora a descoberta de dados (Park; Tosaka, 2010). O monitoramento contínuo das tecnologias utilizadas também é essencial para evitar problemas de obsolescência, garantindo que os dados permaneçam acessíveis e utilizáveis ao longo do tempo.

A qualidade dos metadados é outro pilar para a interoperabilidade e a consistência dos dados em repositórios. Autores como Hemphill *et al.* (2024) e Park e Tosaka (2010) ressaltam a necessidade de acompanhar os conjuntos de dados com metadados rigorosamente detalhados. Paralelamente, ferramentas automatizadas têm sido desenvolvidas para avaliar as dimensões de qualidade (precisão, completude e consistência entre outros) dos dados. A automação das avaliações pode aumentar a eficiência e a reprodutibilidade dos processos de garantia de qualidade (Rajan, 2019). Essa prática não só melhora a organização e acessibilidade dos dados, mas também reduz ambiguidades e erros durante a análise.

A adesão a certificações como o *CoreTrustSeal* atesta o compromisso dos repositórios com padrões elevados de gestão e preservação de dados. Essas certificações validam a implementação de políticas de curadoria, aumentando a confiança da comunidade científica (Kindling; Strecker, 2022; Aerts *et al.*, 2021). Paralelamente, auditorias periódicas e a aplicação de métricas de qualidade

SUMÁRIO

contribuem para a melhoria contínua dos repositórios, garantindo que os dados permaneçam úteis e confiáveis ao longo do tempo.

Assim, a curadoria ativa e a realização de auditorias sistemáticas são práticas indispensáveis para manter a qualidade dos dados ao longo do tempo. A qualidade dos dados em repositórios de pesquisa exige uma combinação de métodos que envolvem a participação humana, a automação de processos e a adoção de padrões reconhecidos internacionalmente.

Outro aspecto destacado na literatura é a criação de declarações de propósito para cada conjunto de dados (Andrews *et al.*, 2024), assegurando sua relevância e alinhamento com os objetivos de pesquisa. Ademais, a padronização dos dados (Huser *et al.*, 2018), frequentemente obtida por meio de modelos comuns, facilita tanto o compartilhamento quanto a análise.

A implementação de modelos de qualidade, como os mencionados por Andrews (2024), Hemphill *et al.* (2024), Rajan (2019) e Huser *et al.* (2018), que exploram sistemas automatizados de conhecimento, reforça a importância de alinhar as dimensões tradicionais de qualidade – precisão, consistência e completude – com as demandas atuais de interoperabilidade e escalabilidade.

Lavery *et al.* (2022) descrevem o processo de garantia de qualidade dos dados na medicina de precisão em oncologia para assegurar a integridade dos dados, incluindo avaliações de viabilidade, precisão e reprodutibilidade. Os autores reforçam a importância da curadoria no processo de verificação de dados para garantir a precisão, destacando a curadoria dupla e revisão de código para avaliar a reprodutibilidade.

Os processos de avaliação da qualidade podem identificar erros e aprimorar métodos de correção por meio do treinamento dos curadores e ajustes nas diretrizes de curadoria. Dessa forma, os dados podem ser considerados de alta qualidade e prontos

SUMÁRIO

para análises que apoiem a tomada de decisões clínicas. Embora a curadoria manual seja demorada, os processos descritos garantem a qualidade dos dados. O processo apresentado por Lavery *et al.* (2022) é escalável e pode ser adaptado para outras iniciativas de curadoria de dados em grande escala.

Já o projeto Quality - *Established: Testing and application of curation criteria and quality standards for audiovisual annotated language data* (QUEST) tem como objetivo estabelecer critérios de qualidade e curadoria para dados linguísticos de anotações audiovisuais (Arkhangelskiy; Hedeland; Riaposov, 2020). O QUEST busca promover formatos e padrões que facilitem o controle de qualidade automático, o que inclui a interoperabilidade semântica entre diferentes formatos e convenções.

Para garantir a qualidade dos dados, o projeto propõe três níveis de maturidade dos dados: dados com metadados mínimos, sem critérios específicos de qualidade; dados com estrutura consistente e descrição das relações entre os objetos; e dados com estrutura e conteúdo consistentes, incluindo anotações e metadados contextuais. Essa divisão permite uma avaliação mais adequada da qualidade dos dados, evitando que critérios rigorosos impeçam o depósito de dados valiosos.

Os autores Arkhangelskiy, Hedeland e Riaposov (2020) relatam que o projeto desenvolveu um sistema de controle de qualidade que inclui um questionário para os responsáveis pelo depósito dos dados. Uma ferramenta *web* ajuda os pesquisadores a planejar projetos e avaliar a qualidade dos dados antes do depósito. O questionário gera modelos de arquivos de transcrição e anotações, garantindo consistência e reduzindo o trabalho de curadoria posterior. O sistema conta ainda com um *framework* de controle de qualidade por meio de uma extensão do *Corpus Services* do HZSK. O *framework* realiza verificações automáticas de qualidade, como validação de XML, cobertura de *links*, estrutura de arquivos e consistência de

SUMÁRIO

anotações. Uma base de dados multilíngue apresenta recomendações e explicações sobre critérios de qualidade, formatos de dados e melhores práticas. A base de conhecimento é acessível a pesquisadores e ajuda a garantir a adesão aos critérios estabelecidos.

Em se tratando de análises de *big data* e aplicações para *internet of things* (IoT), a qualidade dos dados é um fator crítico para a eficácia de sistemas de informação e exige abordagens adaptativas. Reda, Benabdellah e Zellou (2023) realizaram uma revisão sistemática de literatura sobre avaliação da qualidade de dados, abrangendo modelos, metodologias, métricas e domínios de aplicação e identificaram tendências, lacunas e desafios. Os resultados mostram que completude, precisão e consistência são as dimensões mais avaliadas.

Entre os modelos, o estudo aponta a observância dos Princípios FAIR para dados de pesquisa, a abordagem holística para a gestão (*Total Data Quality Management*) e as métricas baseadas em critérios estatísticos (*Information Quality Metrics*) para a avaliação da qualidade dos dados (Reda; Benabdellah; Zellou, 2023). *Frameworks* como o *Data Quality for Smart Data* (DQ2SD) para *big data*; padrões *Fast Healthcare Interoperability Resources* (FHIR) para registros de procedência de dados de prontuários eletrônicos na área da saúde utilizando a lógica *fuzzy* para determinar o nível de confiabilidade dos dados e ferramentas como Hygieia para validação em tempo real para a IoT são abordagens por domínio encontradas no estudo de Reda, Benabdellah e Zellou (2023).

Diante da complexidade e diversidade dos dados de pesquisa, garantir sua qualidade requer mais do que práticas isoladas. Trata-se de um processo contínuo e multifacetado, que deve integrar avaliações sistemáticas, intervenções cuidadosamente planejadas e estratégias eficazes de comunicação. A construção de um ecossistema de dados confiável depende da colaboração entre curadores, pesquisadores, desenvolvedores de ferramentas, gestores de repositórios e usuários finais.

SUMÁRIO

O uso combinado de metodologias automatizadas, padrões reconhecidos internacionalmente e modelos de ciclo de vida consolidados reforça a necessidade de abordagens adaptativas, alinhadas às especificidades de cada domínio científico. Além disso, iniciativas como o projeto QUEST e as aplicações em *big data* e IoT mostram que a qualidade dos dados não é uma preocupação exclusiva da ciência aberta, mas uma demanda transversal que impacta diretamente a inovação tecnológica, a integridade científica e a tomada de decisões informadas. A aplicação de critérios como os princípios FAIR, o uso de *frameworks* especializados e a curadoria ativa demonstram que há um esforço global em consolidar boas práticas e modelos sustentáveis de gestão da qualidade dos dados.

FERRAMENTAS PARA AVALIAÇÃO DA QUALIDADE DOS DADOS (AQD)

Métodos de AQD destacam a necessidade de ferramentas e *frameworks* para garantir a QD dados em pesquisas (Huser *et al.*, 2018). A identificação da QD garante que as análises e modelos gerados sejam baseados em informações confiáveis e precisas. Diversas ferramentas e serviços têm sido desenvolvidos, permitindo que profissionais avaliem dimensões como completude, precisão, consistência, atualidade e acessibilidade dos dados. Estas ferramentas são especialmente úteis em um cenário onde a governança de dados é cada vez mais desafiadora devido à complexidade e ao volume crescente de dados disponíveis.

Entre as soluções mais avançadas de ferramentas, Huser *et al.* (2018) destacam as ferramentas da área da saúde, como a Achilles Heel, desenvolvida pelo consórcio *Observational Data Science and Informatics* (OHDSI), que avalia a conformidade dos dados com o modelo *Observational Medical Outcomes Partnership* (OMOP).

SUMÁRIO

A ferramenta incorpora uma tabela de metadados para capturar descrições de dados estruturados e não estruturados para melhorar a interpretação dos dados. A consistência e a integridade de dados em repositórios são testadas, assegurando que os padrões estabelecidos sejam seguidos.

Outra abordagem apresentada por Huser *et al.* (2018) é a do PEDSnet que desenvolveu um extenso programa de avaliação de qualidade de dados. Com mais de 11.000 regras de qualidade e convenções detalhadas, o PEDSnet resolveu discrepâncias causadas por diferentes interpretações na harmonização de dados.

O modelo de revisão em quatro níveis, adotada pela Iniciativa Sentinel e apoiada pela *Food and Drug Administration* (FDA), dos Estados Unidos, utiliza o *Sentinel Common Data Model* (SCDM), para a verificação da conformidade, completude, validação e integridade entre as tabelas de dados (Huser *et al.*, 2018).

Para enfrentar a heterogeneidade dos dados, Huser *et al.* (2018) salientam que a arquitetura *Open Quality and Analytics Framework* (OQAF) surge como uma solução baseada em três componentes principais: a) um repositório de conhecimento em qualidade, *Quality Knowledge Repository* (QKR), para desenvolver um metamodelo de QD baseado em um repositório de metadados padrão que armazena artefatos de metadados e seus relacionamentos; b) uma plataforma de integração federada utilizando a plataforma *OpenFurther* (OF), que suporta consultas centradas em metadados semanticamente consistentes de fontes de dados heterogêneas para pesquisa translacional usando um repositório de metadados e c) um *meta-framework* de visualização, *Visualization Meta-Framework* (VMF), que fornece um repositório para armazenar visualizações para diferentes AQD e suas métricas. A arquitetura permite avaliações padronizadas e agnósticas de modelos de dados específicos.

SUMÁRIO

Percebe-se que diversas ferramentas automatizadas têm sido adotadas para avaliar dimensões críticas da QD. Rajan (2019) enfatiza que a aplicação de *frameworks* de qualidade de dados e metodologias de avaliação em domínios como Qualidade de Dados e Qualidade da Informação, automatizam a avaliação da qualidade em ambientes heterogêneos.

Encontram-se ainda, soluções como *OpenRefine*, *FAIR Evaluators* e *DQMiner* que permitem análises rápidas e reproduzíveis, reduzindo erros humanos e aumentando a eficiência dos processos de curadoria. Tem-se utilizado a aplicação dos princípios FAIR como estratégia para melhorar a qualidade dos dados. A ferramenta *5-Star Data Rating Tool* permite verificar a conformidade dos repositórios com os princípios FAIR (Rodrigues; Dias; Lourenço. 2022).

Hemphill *et al.* (2024) apontam que a ferramenta Jira é utilizada para rastrear e gerenciar tarefas relacionadas à curadoria e manutenção de conjuntos de dados, facilitando o monitoramento e a qualidade dos dados.

Em uma abordagem anterior, Visuri *et al.* (2019) preocupam-se com o controle de QD em um projeto piloto para a criação de um repositório de dados. O processo de controle de qualidade relatado pelos autores é executado com um módulo automatizado *QualityGuard*, e funciona como parte do processo de importação para o banco de dados. Os produtores de dados também podem usá-lo de forma independente para encontrar possíveis erros em seus dados. O *QualityGuard* retorna um relatório de QD, que contém estatísticas e um conjunto de informações de erro de qualidade referente aos dados importados.

Ressaltando o processo multifacetado da AQD em repositórios de dados de pesquisa, o estudo realizado por Kindling e Strecker (2022) apresenta as principais ferramentas e práticas identificadas. A avaliação formal dos dados concentra-se em aspectos técnicos,

SUMÁRIO

administrativos e de acesso. Entre as ferramentas e práticas mais comuns estão a verificação de metadados, os autores destacam: a) observância de uma descrição básica dos dados, utilizando esquemas de metadados padronizados, como o *Data Documentation Initiative* (DDI) ou *Dublin Core*; b) conformidade com os Princípios FAIR, utilizando ferramentas como o *FAIR Evaluator* ou *FAIRshake*; c) verificação de direitos e licenças para garantir a conformidade legal; e d) testes de integridade física com ferramentas automatizadas, como *checksums* ou algoritmos de verificação de arquivos para garantir que os dados não foram corrompidos durante o armazenamento ou transferência.

Palepu e Rao (2012) destacam a importância de abordar problemas de qualidade desde a fase inicial de construção de um repositório e propõem um modelo baseado em metadados para mitigar anomalias. A falta de validação e inconsistências na formatação da heterogeneidade semântica das fontes de dados, as falhas na aplicação de regras de negócio ou na atualização periódica dos dados, bem como deficiências na análise de requisitos ou escolha inadequada de modelos dimensionais na modelagem de esquema produzem desafios de qualidade nas múltiplas etapas da criação de um repositório de dados. Os autores categorizam as ferramentas de avaliação em três grupos: auditoria, limpeza e migração. A auditoria verifica a conformidade com regras de negócio e identifica discrepâncias; a limpeza padroniza e valida dados; e a migração, que envolve a transferência e transformação dos dados entre plataformas, assegura a compatibilidade. Além disso, padrões como acurácia, integridade e consistência são essenciais para métricas de qualidade, alinhados a regulamentações e necessidades do negócio.

Assim, Palepu e Rao (2012) delineiam uma arquitetura de controle de qualidade baseada em metadados. A proposta integra a transformação em especificações técnicas de requisitos do usuário, o monitoramento contínuo com a avaliação quantitativa ao longo do

SUMÁRIO

fluxo de dados e um repositório de metadados que armazena características dos dados para identificar e corrigir problemas proativamente.

Kapsner *et al.* (2021) detalham o desenvolvimento de uma ferramenta de avaliação de qualidade de dados para repositórios de dados de pesquisa no âmbito do consórcio MIRACUM. A ferramenta foi projetada para avaliar a qualidade dos dados de registros eletrônicos de saúde usados em estudos observacionais, assegurando que os dados sejam adequados para análises locais e federadas. São avaliadas a conformidade, a completude e a plausibilidade dos dados, seguindo um *framework* harmonizado proposto por Kahn *et al.* (2016).

É importante destacar a importância da qualidade dos dados em cenários onde os dados são reaproveitados, ou seja, usados para fins diferentes daqueles para os quais foram originalmente criados. Os autores Zhang, Indulska e Sadiq (2019) ressaltam que as metodologias de avaliação de qualidade de dados são centradas no usuário e seguem uma abordagem *top-down*. Os usuários frequentemente se deparam com conjuntos de dados desconhecidos, sobre os quais não têm controle ou conhecimento prévio. Isso torna inviável a aplicação de abordagens tradicionais de avaliação de qualidade.

O trabalho de Zhang, Indulska e Sadiq (2019) propõe uma abordagem, chamada LANG, que permite a descoberta de problemas de qualidade de dados de forma sistemática e repetível, mesmo em cenários onde a criação e o uso dos dados são separados. A abordagem LANG observa as dimensões de completude, precisão, consistência e conformidade.

A abordagem LANG é composta por duas etapas principais, a etapa sintática, com foco na identificação de problemas relacionados à conformidade dos dados com regras sintáticas, como unicidade, consistência de formato, integridade referencial e conformidade com metadados; e a etapa semântica, a qual concentra-se na detecção de problemas relacionados ao significado dos dados, como completude

SUMÁRIO

de atributos obrigatórios e opcionais, consistência semântica, consistência de valores e precisão (Zhang; Indulska; Sadiq, 2019). A abordagem foi validada empiricamente em termos de solidez, repetibilidade e generalização, utilizando conjuntos de dados reais.

A referida abordagem preenche uma lacuna importante na literatura e na prática, fornecendo uma metodologia sistemática para a descoberta de problemas de qualidade de dados em conjuntos de dados reaproveitados. Particularmente útil em cenários onde os usuários não têm controle sobre a criação dos dados ou acesso a metadados detalhados.

Outra abordagem de avaliação da qualidade é a *Data Integration Unit* (DIU) utilizada para o armazenamento de dados a longo prazo em repositórios de dados controlados pelo custodiante, abordada por Schneider *et al.* (2018). O DIU realiza atividades de controle de qualidade. Ciclos de *feedback* incorporados ao processo de qualidade dos dados permitem que os custodiantes melhorem seus conjuntos de dados com base em problemas identificados pelo DIU. O controle de qualidade extensivo realizado pelo DIU garante que a qualidade dos dados fornecidos aos pesquisadores seja aprimorada.

Como demonstrado, ferramentas e *frameworks* oferecem soluções especializadas para dimensionar e aprimorar a completude, precisão, consistência e conformidade dos dados. A adoção de *frameworks* padronizados, dos princípios FAIR e a integração de processos automatizados de controle de qualidade reforçam a importância de uma gestão proativa dos dados, minimizando erros e maximizando sua reutilização. Além disso, a evolução contínua dessas tecnologias e a criação de novas ferramentas evidenciam a necessidade de abordagens adaptativas, capazes de lidar com a heterogeneidade e o volume crescente de dados. A colaboração entre domínios, a harmonização de padrões e o investimento em metadados robustos são passos essenciais para assegurar que os dados não apenas atendam a critérios técnicos rigorosos, mas também sustentem análises confiáveis e decisões informadas.

CONSIDERAÇÕES FINAIS

A qualidade dos dados tem sido uma área de pesquisa por mais de duas décadas, com contribuições da ciência da computação, estatística, sistemas de informação e áreas de distintos domínios. Há dois momentos importantes na vida útil dos dados: o momento em que ele é criado e o momento em que ele é usado. O gerenciamento de qualidade de dados se refere a conectar esses momentos no tempo para garantir que o momento da criação seja executado e gerenciado corretamente, para que tudo corra bem no momento do uso.

A avaliação da qualidade dos dados de pesquisa manifesta-se como um pilar fundamental para a integridade, confiabilidade e reprodutibilidade das pesquisas científicas. A QD não se resume apenas à precisão, mas abrange múltiplas dimensões, como completude, consistência, acessibilidade e relevância, que variam conforme o contexto de uso dos dados. A crescente dependência de dados em ambientes acadêmicos, empresariais e tecnológicos exige abordagens robustas para garantir que as informações sejam adequadas aos seus propósitos, evitando vieses, inconsistências e perda de valor.

Os repositórios de dados de pesquisa (RDP) desempenham um papel central nesse cenário, atuando na preservação da autenticidade, integridade e privacidade dos dados. A adoção de princípios como os FAIR (Localizáveis, Acessíveis, Interoperáveis e Reutilizáveis) e a implementação de ferramentas automatizadas, como *OpenRefine* e *FAIR evaluators*, entre outras citadas neste estudo, têm se mostrado eficazes para aprimorar a qualidade e facilitar o compartilhamento e reuso dos dados. Além disso, a curadoria ativa, a governança de dados e a adesão a padrões internacionais, como ISO 8000 e ISO 9000, são estratégias essenciais para manter a confiabilidade dos dados ao longo do tempo.

SUMÁRIO

No entanto, os desafios persistem. A heterogeneidade dos dados, a falta de padronização entre disciplinas e a necessidade de conscientização dos pesquisadores sobre práticas de gestão de dados destacam-se como obstáculos a serem superados. A “crise de reprodutibilidade” evidenciada em estudos recentes reforça a urgência de políticas institucionais e culturais que incentivem a transparência, a documentação detalhada e o acesso aberto aos dados e métodos.

Por fim, este estudo destaca que a qualidade dos dados é um processo contínuo e colaborativo, que envolve pesquisadores, curadores, desenvolvedores de ferramentas e gestores de repositórios. Investir em educação sobre boas práticas, desenvolver *frameworks* adaptáveis e promover a interoperabilidade entre sistemas são passos primordiais para construir um ecossistema de dados confiável e sustentável. A qualidade dos dados não é apenas uma questão técnica, mas uma condição indispensável para o avanço da ciência, a tomada de decisão e a inovação responsável.

Assim, conclui-se que a busca por dados de alta qualidade deve ser uma prioridade coletiva, alinhada aos princípios da Ciência Aberta e às demandas de um mundo cada vez mais orientado por dados. A implementação das estratégias e ferramentas discutidas neste trabalho pode pavimentar o caminho para pesquisas mais robustas, transparentes e impactantes, beneficiando não apenas a comunidade científica, mas a sociedade como um todo.

REFERÊNCIAS

ABEDJAN, Z.; GOLAB, L.; NAUMANN, F. Data quality tools for big data: a comprehensive review and comparison. **Journal of Big Data**, Berlin, v. 3, n. 1, p. 1-18, 2016.

AERTS, H. *et al.* Quality of hospital electronic health record (EHR) data based on the International Consortium for Health Outcomes Measurement (ICHOM) in heart failure: pilot data quality assessment study. **JMIR Medical Informatics**, Toronto, v. 9, n. 8, e27842, 2021. DOI: <https://doi.org/10.2196/27842>.

SUMÁRIO

ANDREWS, J. T. A. *et al.* Ethical considerations for responsible data curation. *In: Proceedings of the 37th International Conference on Neural Information Processing Systems (NIPS '23)*. Red Hook, NY, USA: Curran Associates Inc., 2024. Art. 2415, p. 55320-55360.

ARAÚJO, D. G.; DIAS, G. A.; AUTRAN, M. M. M. Compartilhamento de dados no contexto da ciência brasileira: um estudo integrativo. **Informação & Informação**, Londrina, v. 26, n. 3, p. 100-124, jul./set. 2021. DOI: <https://doi.org/10.5433/1981-8920.2021v26n3p100>.

ARKHANGELSKIY, T.; HEDELAND, H.; RIAPOSOV, A. Evaluating and assuring research data quality for audiovisual annotated language data. **CLARIN Annual Conference**, 2020.

ASSANTE, M. *et al.* Are Scientific Data Repositories Coping with Research Data Publishing? **Data Science Journal**, Cambridge, v. 15, p. 1-24, 2016. Disponível em: <https://digitalcommons.unl.edu/scholcom/311/>. Acesso em: 05 abr. 2025.

BAKER, M. 1,500 scientists lift the lid on reproducibility. **Nature**, London, v. 533, p. 452-454, maio 2016. DOI: <https://doi.org/10.1038/533452a>.

BARBA, L. A. Terminologies for Reproducible Research. **Computer Science: Digital Libraries**, Berlin, 2018. DOI: <https://doi.org/10.48550/arXiv.1802.03311>.

BERTOSSI, L.; RIZZOLO, F. Contexts and Data Quality Assessment. **ArXiv**, Ithaca, 2016. DOI: <https://doi.org/10.48550/arXiv.1608.04142>.

BOYD, D.; CRAWFORD, K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, Londres, v. 15, n.5, p. 662-679, 2012. Disponível em: DOI: <https://doi.org/10.1080/1369118X.2012.678878>. Acesso em: 05 abr. 2025

CAREGNATO, S. E.; ROCHA, R. P. da; Rene; F. G. J. Reúso de dados: princípios FAIR e o ecossistema de pesquisa. *In: SALES, L. F.; VEIGA, V. dos S.; HENNING, P.; SAYÃO, L. F. (org.). Princípios FAIR aplicados à gestão de dados de pesquisa*. Rio de Janeiro: Ibict, 2021. p. 187 - 200. Disponível em: https://www.researchgate.net/publication/355672623_Reuso_de_dados_principios_FAIR_e_o_ecossistema_de_pesquisa. Acesso em: 6 abr. 2025.in:

CHEN, X. *et al.* Open is not enough. **Nature Physics**, London, v. 15, p. 113-119, fev./2019. Disponível em: <https://www.nature.com/articles/s41567-018-0342-2>. Acesso em: 8 dez. 2024.

SUMÁRIO

DAVENPORT, T. H.; PATIL, D. J. Data scientist: The sexiest job of the 21st century. **Harvard Business Review**, Brighton, v. 90, n. 10, p. 70-76, 2012. Disponível em: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>. Acesso em: 13 abr. 2025.

FEINBERG, M. *et al.* The New Reality of Reproducibility: The Role of Data Work in Scientific Research. **Proceedings of the ACM on Human-Computer Interaction**, New York, v. 4, n. 1, p. 1-22, maio/2020. Disponível em: <https://doi.org/10.1145/3392840>. Acesso em: 05 abr. 2025.

FERREIRA, M. K. **Dados de pesquisa**: contribuição da infraestrutura para a promoção da reprodutibilidade e do reúso. 2024. 182 f. Tese (Doutorado em Comunicação) - Programa de Pós-Graduação em Comunicação, Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2024. Disponível em: <https://lume.ufrgs.br/handle/10183/275864>. Acesso em: 05 abr. 2025.

FOUAD, K *et al.* A practical guide to data management and sharing for biomedical laboratory researchers. **Experimental Neurology**, Amsterdam, v. 378, p. 1-14, maio/2024. DOI: <https://doi.org/10.1016/j.expneurol.2024.114815>.

GOBLE, C.; RURE, D., BECHHOFFER, S. Accelerating Scientists' Knowledge Turns. In: FRED, A.; DIETZ, J.L.G., LIU, K.; FILIPE, J. (orgs.) **Knowledge Discovery, Knowledge Engineering and Knowledge Management**. IC3K 2011. Communications in Computer and Information Science, Berlin, v. 348, p. 25, 2013. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-37186-8_1.

HARRIS, J. K. *et al.* Use of reproducible research practices in public health: A survey of public health analysts. **PLoS ONE**, San Francisco, v. 13, n. 9, p. e0202447, 12 set. 2018. DOI: <https://doi.org/10.1371/journal.pone.0202447>. Acesso em: 05 abr. 2025.

GOODMAN, S. N.; FANELLI, D.; IOANNIDIS, J. P. A. What does research reproducibility mean? **Science Translational Medicine**, Washington, v. 8, n. 341, p. 14, jun. 2016. Disponível em: <https://www.science.org/doi/10.1126/scitranslmed.aaf5027>. Acesso em: 8 dez. 2025.

HEMPHILL, L. *et al.* A dataset for measuring the impact of research data and their curation. **Scientific Data**, London, v. 11, n. 1, p. 442, 2024. DOI: <https://doi.org/10.1038/s41597-024-03303-2>.

SUMÁRIO

HUSER, V. *et al.* Methods for examining data quality in healthcare integrated data repositories. *In: Pacific Symposium on Biocomputing*. 23. Honolulu, Hawaii: Pacific Symposium on Biocomputing, p. 628-633, 2018.

HJØRLAND, B. **Data**: with Big Data and database semantics. *In: ISKO Encyclopedia of Knowledge Organization*. 2018. Disponível em: <http://www.isko.org/cyclo/data>. Acesso em: 04 jun. 2023.

KAHN, M. G. *et al.* A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. **EGEMS**, London, v. 4, n. 1, 2016.

KANDEL, S. *et al.* Research directions in data wrangling: Visualizations and transformations for usable and credible data. **Information Visualization**, London, v. 10, n. 4, 271-288, 2011. DOI: <https://doi.org/10.1177/1473871611415994>.

KAPSNER, L. A. *et al.* Linking a consortium-wide data quality assessment tool with the MIRACUM metadata repository. **Applied Clinical Informatics**, Stuttgart, v. 12, n. 4, p. 826-835, 2021. DOI: <https://doi.org/10.1055/s-0041-1733847>.

KIM, Y.; ZHANG, P. Understanding data sharing behaviors of STEM researchers: The roles of attitudes, norms, and data repositories. **Library & Information Science Research**, Amsterdam, v. 37, n. 3, p. 189-200, jul./2015. DOI: <https://doi.org/10.1016/j.lisr.2015.04.006>.

KINDLING, M.; STRECKER, D. Data Quality Assurance at Research Data Repositories. **Data Science Journal**, Paris, v. 21, n. 18, pp. 1-17, 2022. DOI: <https://doi.org/10.5334/dsj-2022-018>.

KOLTAY, T. Data literacy: in search of a name and identity. **Journal of Documentation**, Bingley, West Yorkshire, v. 71, n. 2, p. 401-415, 2015. DOI: <https://doi.org/10.1108/JD-02-2014-0026>. Acesso em: 23 nov. 2024.

KOLTAY, T. Data literacy for researchers and data librarians. **Journal of Librarianship and Information Science**, London, v. 49, n. 1, p. 3-14. 2017. Disponível em: <https://journals.sagepub.com/doi/abs/10.1177/0961000615616450>. Acesso em: 22 jan. 2023.

LAVERY, J. A. *et al.* A scalable quality assurance process for curating oncology electronic health records: the project genie biopharma collaborative approach. **JCO Clinical Cancer Informatics (JCO CCI)**, Alexandria, v. 6, 2022.

MCKIERNAN, E. C. *et al.* How open science helps researchers succeed. **eLife**, Cambridge, v. 5, p. 1-19, 7 jul. 2016. DOI: <https://doi.org/10.7554/eLife.16800>.

SUMÁRIO

MONTEIRO, E. C. de S. de A. **Operacionalização de repositórios de dados:** uma análise sobre as perspectivas e atitudes dos pesquisadores nas questões de autoria e licença. 272 f. 2021. (Tese - Doutorado em Ciência da Informação) - Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília. Disponível em: <http://hdl.handle.net/11449/214671>. Acesso em: 05 abr. 2025.

MONTEIRO, E. C. de S. de A.; SANT'ANA, R. C. G. Anomia de pesquisadores no compartilhamento de dados. **Em Questão**, Porto Alegre, v. 29, p. 1-24, 2023. DOI <https://doi.org/10.19132/1808-5245.29.122627>.

MONTEIRO, E. C. de S. de A.; SANT'ANA, R. C. G. Cultura e anomia no compartilhamento de dados: uma perspectiva a partir dos pesquisadores. **Transinformação**, Campinas, v. 36, p. 1-13, 2024. DOI: <https://doi.org/10.1590/2318-0889202436e247095>.

MORTON, L. 5 Open Science practices that improve reproducibility & support trust in science. **PLOS Blog**, San Francisco, 14 jul. 2022. Disponível em: <https://theplosblog.plos.org/2022/07/reproducibility/>. Acesso em: 08 dez. 2025.

OLSON, J. E. **Data Quality: The Accuracy Dimension**. San Francisco: Morgan Kaufmann, 2003. (The Morgan Kaufmann Series in Data Management Systems).

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. OECD Principles and Guidelines for Access to Research Data from Public Funding. OECD, 2007. DOI 10.1787/9789264034020-en-fr. Disponível em: <https://www.oecd.org/sti/inno/38500813.pdf>. Acesso em: 22 jan. 2023.

PAGANO, T. P. *et al.* Bias and unfairness in machine learning models: a systematic review on datasets, tools, fairness metrics, and identification and mitigation methods. **Big Data Cognitive Computing**, Basel, 2023, v. 7, n. 15. Disponível em: <https://www.mdpi.com/2504-2289/7/1/15>. Acesso em: 18 nov. 2024.

PALEPU, R. B.; RAO, K. V. S. Meta data quality control architecture in data warehousing. **International Journal of Computer Science, Engineering and Information Technology (IJCEIT)**, Chennai, v.2, n. 4, 2012.

PARK, J.-R.; TOSAKA, Y. Metadata quality control in digital repositories and collections: criteria, semantics, and mechanisms. **Cataloging & Classification Quarterly**, Philadelphia, v. 48, n. 8, p. 696-715, 2010. DOI: <http://dx.doi.org/10.1080/01639374.2010.508711>.

PENG, R. D.; HICKS, S. C. **Reproducible Research: a Retrospective**. Annual Review of Public Health, 42: 2021. p. 79-93. DOI: <https://doi.org/10.1146/annurev-publhealth-012420-105110>.

SUMÁRIO

PETTIBONE, K. G. *et al.* Expanding the Concept of Translational Research: Making a Place for Environmental Health Sciences. **Environmental Health Perspectives**, v. 126, n. 7, Jul. 2018. DOI: <https://doi.org/10.1289/EHP3657>.

PIMENTEL, J. F. *et al.* A large-scale study about quality and reproducibility of Jupyter Notebooks. *In*: IEEE/ACM INTERNATIONAL CONFERENCE ON MINING SOFTWARE REPOSITORIES (MSR), 16., 2019, **Proceedings**. Los Alamitos: IEEE Computer Society, 2019. p. 507-517. DOI: <https://doi.org/10.1109/MSR.2019.00077>.

PIWOWAR, H. A.; VISION, T. J. Data reuse and the open data citation advantage. **PeerJ**, San Diego, v. 1, p. e175, 2013. DOI: <https://doi.org/10.7717/peerj.175>.

PRESS, G. Cleaning Big Data: most time-consuming, least enjoyable data science task, survey says. **Forbes**, 2016. Disponível em: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>. Acesso em: 18 nov. 2024.

PROVOST, F.; FAWCETT, T. **Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking**. Sebastopol: O'Reilly Media, 2013. Capítulo 1. Disponível em: <https://www.amazon.com/Data-Science-Business-Data-Analytic-Thinking/dp/1449361323?asin=1449361323&revisionId=&format=4&depth=1>. Acesso em: 13 abr. 2025.

RAHM, E.; DO, H. H. Data cleaning: Problems and current approaches. **IEEE Data Engineering Bulletin**, Washington, v. 23, n. 4, p. 3-13, 2020. Disponível em: https://www.researchgate.net/publication/220282831_Data_Cleaning_Problems_and_Current_Approaches. Acesso em: 13 abr. 2025.

RAJAN, N. S. *et al.* Towards a content agnostic computable knowledge repository for data quality assessment. **Computer Methods and Programs in Biomedicine**, Amsterdam, v. 177, p. 193-201, 2019. DOI: <https://doi.org/10.1016/j.cmpb.2019.05.017>.

GERMAN COUNCIL FOR SCIENTIFIC INFORMATION INFRASTRUCTURES. **The Data Quality Challenge**: Recommendations for Sustainable Research in the Digital Turn, Göttingen: RFI, 2020, 120 p.

REDA, O.; BENABDELLAH, N. C.; ZELLOU, A. A systematic literature review on data quality assessment. **Bulletin of Electrical Engineering and Informatics**, Yogyakarta, v. 12, n. 6, p. 3736-3757, dez. 2023. ISSN 2302-9285. DOI 10.11591/eei.v12i6.5667.

SUMÁRIO

REIS, M. J. **Ciência da Informação e Ciência de Dados**: guia para alfabetização de dados para bibliotecários. 2019. Dissertação (Mestrado Profissional em Gestão da Informação e do Conhecimento) – Universidade Federal de Sergipe. 143 f. 2019. Disponível em: https://ri.ufs.br/bitstream/riufs/12667/2/MAKSON_%20JESUS_REIS.pdf. Acesso em: 22 nov. 2024.

RESEARCH DATA OXFORD. Sharing data. 2024. Disponível em: <https://researchdata.ox.ac.uk/sharing-data>. Acesso em: 08 dez. 2025.

RODRIGUES, M. M.; DIAS, G. A.; LOURENÇO, C. de A. Repositórios de dados científicos na América do Sul: uma análise da conformidade com os Princípios FAIR. **Em Questão**, Porto Alegre, v. 28, n. 2, p. 113057, 2022. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/113057>. Acesso em: 13 out.

SADIQ, S.; INDULSKA, M. Open data: Quality over quantity. **International Journal of Information Management**, Amsterdam, v. 37, n. 3, p. 150-154, 2017. DOI: <https://doi.org/10.1016/j.ijinfomgt.201701.003>.

SALOMÃO, A. **Reproducibility vs. Replicability**: Key Differences. MIND THE GRAPH. 2023. Disponível em: <https://mindthegraph.com/blog/reproducibility/#:~:text=Both%20reproducibility%20and%20replicability%20are,range%20of%20situations%20or%20contexts>. Acesso em: 8 dez. 2024.

SAMUEL, S.; KÖNIG-RIES, B. Understanding experiments and research practices for reproducibility: an exploratory study. **PeerJ**, San Diego, v. 9, p. 1-26, 2021. DOI: <https://doi.org/10.7717/peerj.11140>.

SANCHEZ, F. A.; VIDOTTI, S. A. B. G.; VECHIATO, F. L. A contribuição da curadoria digital em repositórios digitais. **Revista Informação na Sociedade Contemporânea**, Natal, v. 1, p. 1-17, 2017. Disponível em: <https://periodicos.ufrn.br/informacao/article/view/12280>. Acesso em: 13 out. 2024.

SANT'ANNA, J.; DIAS, C. C.; MACULAN, B. C. A gestão dos dados de pesquisa nas universidades e o papel dos serviços informacionais oferecidos nas bibliotecas: uma revisão narrativa. **Múltiplos olhares em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, p. 1-16, 2019. Disponível em: <https://periodicos.ufmg.br/index.php/moci/article/view/19169/16262>. Acesso em: 22 nov. 2024.

SUMÁRIO

SANTOS, P. X.; *et al.* **Livro Verde - Ciência aberta e dados abertos**: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz, 2017. ISBN: 978-85- 8110-036- 4. Disponível em: <https://www.arca.fiocruz.br/handle/icict/24117>. Acesso em: 20 nov. 2024.

SAYÃO, L. F.; SALES, L. Afinal, o que é dado de pesquisa? **Biblos**, Rio Grande, v. 34, n. 02, p. 32-51, jul./dez. 2020. DOI 10.14295/biblos.v34i2.11875. Disponível em: <https://periodicos.furg.br/biblos/article/view/11875>. Acesso em: 22 jan. 2023.

SAYÃO, L. F.; SALES, L. **Guia de Gestão de Dados de Pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/ IEN, 2015. Disponível em: https://www.researchgate.net/publication/286455028_Guia_de_gestao_de_dados_de_pesquisa_para_bibliotecarios_e_pesquisadores. Acesso em: 6 abr. 2025.

SCHNEIDER, M. *et al.* Supporting the timely, easy and cost effective access to high-quality linked data via the Custodian-Controlled Data Repository. **International Journal of Population Data Science**, Swansea, v. 3, n. 3, p. 100, 2018.

SEMLER, A. R.; PINTO, A. L. Os diferentes conceitos de dados de pesquisa na abordagem da biblioteconomia de dados. **Ciência da Informação**, Brasília, DF, v. 48, n. 1, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/111525>. Acesso em: 22 jan. 2023.

SILVA, L. C.; PRADO, C. B.; ARAÚJO, E. A. Fatores que influenciam o processo de compartilhamento e reuso de dados de pesquisa: revisão sistemática de literatura a partir da base de dados BRAPCI. **AtoZ: novas práticas em informação e conhecimento**, Curitiba, v. 13, p. 1-15, 2024. DOI: 10.5380/atoz.v13i0.91352. Disponível em: <https://revistas.ufpr.br/atoz/article/view/91352>. Acesso em: 6 abr. 2025.

SILVEIRA, J. I. **Acesso aberto a dados de pesquisa em repositórios universitários internacionais**: um estudo sobre políticas de depósito, acesso e uso. 2021. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal do Rio Grande do Sul. 145 p. Disponível em: <https://lume.ufrgs.br/handle/10183/222095>. Acesso em: 22 jan. 2023.

SILVOLA, R. *et al.* Data quality assessment and improvement. **International Journal of Business Information Systems**, Suíça, v. 22, n. 1, p. 62-81, 2016. DOI: <http://dx.doi.org/10.1504/IJBIS.2016.075718>.

STVILIA, B. An integrated framework for online news quality assurance. **First Monday**, Chicago, v. 26, n. 7, 2021. DOI: <https://doi.org/10.5210/fm.v26i7.11062>.

SUMÁRIO

STVILIA, B. *et al.* A framework for information quality assessment. **Journal of the American Society for Information Science and Technology**, Silver Spring, v. 58, n. 12, p. 1720-1733, 2007.

STVILIA, B.; GIBRADZE, L. Seeking and sharing datasets in an online community of data enthusiasts. **Library and Information Science Research**, Amsterdam, v. 44, p. 101160, 2022.

STVILIA, B.; LEE, D. J. Data quality assurance in research data repositories: a theory-guided exploration and model. **Journal of Documentation**, Bingley, v. 80, n. 4, p. 793-812, 2024. DOI: <https://doi.org/10.1108/JD-09-2023-0177>.

STVILIA, B.; LEE, D. J.; GIBRADZE, L. What is data quality to research data repositories? *In*: ICONFERENCE Bracelona, 2023. **Proceedings** [...]. DOI: 10.13140/RG.2.2.31510.01609.

STVILIA, B. *et al.* Data quality assurance practices in research data repositories: A systematic literature review. An Annual Review of Information Science and Technology (ARIST) paper. **Journal of the Association for Information Science and Technology**, Hoboken, v. 76, n. 1, p. 238-261, 2025. DOI: <https://doi.org/10.1002/asi.24948>.

SIXSIGMA.US. **A Complete guide to data quality management**. 2023. Disponível em: <https://www.6sigma.us/quality/dqm/>. Acesso em: 20 nov. 2024.

STODDEN, V.; SEILER, J.; MA, Z. An empirical analysis of journal policy effectiveness for computational reproducibility. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, DC, v. 115, n. 11, p. 2584-2589, 13 mar. 2018. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.1708290115> Acesso em: 05 abr. 2025.

STODDEN, V.; MIGUEZ, S. Best practices for computational science: software infrastructure and environments for reproducible and extensible research. **SSRN**, p. 1-9, set./2013. Disponível em: <https://ssrn.com/abstract=2322276>. Acesso em: 5 abr. 2025.

STRONG, D. M.; LEE, Y. W.; WANG, R. Y. Data quality in context. **Communications of the ACM**, New York, v. 40, n. 5, p. 103-110, 1997. Disponível em: <https://dl.acm.org/doi/10.1145/253769.253804>. Acesso em: 18 nov. 2024.

TENOPIR, C. *et al.* Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. **PLOS ONE**, San Francisco, v. 15, n. 3, e0229003, 2020. DOI 10.1371/journal.pone.0229003.

TRISOVIC, A *et al.* Repository Approaches to Improving the Quality of Shared Data and Code. **Data**, Basel, v. 6, n. 2, p. 1-12, 2021. DOI: <https://doi.org/10.3390/data6020015>.

SUMÁRIO

VISURI, H. *et al.* Producing and visualizing a country-wide 3d data repository in Finland. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Archives)**, Hanover, v. 42, p. 105–110, 2019. DOI: <https://doi.org/10.5194/isprs-archives-XLII-4-W15-105-2019>.

WAND, Y.; WANG, R. Y. Anchoring data quality dimensions in ontological foundations. **Communications of the ACM**, New York, v. 39, n. 11, p. 86–95, 1996. DOI: <https://doi.org/10.1145/240455.240479>.

WANG, R. Y.; STRONG, D. M. Beyond accuracy: What data quality means to data consumers. **Journal of Management Information Systems**, New York, v. 12, n. 4, p. 5–34, 1996. DOI: <https://doi.org/10.1080/07421222.1996.11518099>.

WICKHAM, H. Tidy data. **Journal of Statistical Software**, California, v. 59, n. 10, p. 1–23, 2014. DOI: <https://doi.org/10.18637/jss.v059.i10>.

WILKINSON, M. D. *et al.* Comments: The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, London, Mar. 2016. DOI: <https://doi.org/10.1038/sdata.2016.18>.

WITT, M. *et al.* Constructing Data Curation Profiles. **International Journal of Digital Curation (IJDC)**, Edinburgh, v. 4, p. 93–103, 2009.

ZHANG, R.; INDULSKA, M.; SADIQ, S. Discovering data quality problems: the case of repurposed data. **Business & Information Systems Engineering (BISE)**, Wiesbaden, v. 61, n. 5, p. 575–593, 2019. DOI: <https://doi.org/10.1007/s12599-019-00608-0>.

13

*Moisés Rockembach
Priscilla Van Even
Jonas Ferrigolo Melo
Luciana Monteiro-Krebs*

METODOLOGIAS DE CO-DESIGN EM CIÊNCIA DA INFORMAÇÃO E COMUNICAÇÃO

DOI: 10.31560/pimentacultural/978-85-7221-475-9.13

SUMÁRIO

COMO CITAR:

ROCKEMBACH, Moisés; VAN EVEN, Priscilla; MELO, Jonas Ferrigolo; MONTEIRO-KREBS, Luciana. Metodologias de Co-design em Ciência da Informação e Comunicação. *In*: SILVA, Fabiano Couto Corrêa da; BARROS, Thiago Henrique Bragato (org.). **Caminhos Metodológicos em Ciência da Informação**: Fundamentos e Aplicações. São Paulo: Pimenta Cultural, 2025.

INTRODUÇÃO

O co-design é um método de pesquisa colaborativa que envolve diversas partes interessadas na criação de soluções, fomentando a participação ativa de sujeitos, usuários finais, especialistas e outras partes envolvidas. Esse método permite que os indivíduos contribuam com suas perspectivas e conhecimentos ao longo de algumas ou todas as etapas do desenvolvimento da pesquisa. Como veremos a seguir, as principais características do codesign incluem a cocriação, a inclusão de múltiplas vozes e a experimentação iterativa, tornando o processo mais democrático e eficaz na resolução de problemas complexos.

Metodologias em Pesquisa em Codesign referem-se a abordagens de design colaborativo que envolvem ativamente diversas partes interessadas na criação de soluções, incluindo usuários finais, especialistas e membros da comunidade. Essa prática evoluiu significativamente ao longo do último século, notavelmente influenciada pelas tradições históricas de design participativo na Escandinávia (Matias, 2011) e por estruturas de cocriação mais recentes introduzidas em diversas organizações. O co-design é reconhecido por seu potencial para aprimorar as experiências do usuário e gerar resultados mais relevantes, integrando as ideias e as experiências vividas pelos participantes ao longo do processo de design.

SUMÁRIO

A relevância das metodologias de co-design reside em sua capacidade de desafiar as dinâmicas tradicionais de poder dentro do processo de design. Ao promover a participação equitativa, essas metodologias priorizam a inclusão, o empoderamento e a propriedade compartilhada entre todas as partes interessadas (Burkett, 2021). Essa abordagem colaborativa não apenas enriquece os resultados do design, mas também promove uma compreensão mais profunda das necessidades e aspirações dos usuários, tornando-a aplicável em diversos setores, como saúde, planejamento urbano, educação e até mesmo bibliotecas (Cestari *et al.*, 2023; Webb *et al.*, 2018; Kirk *et al.*, 2021; Sanders & Stappers, 2008; Abrassart *et al.*, 2015; Bird *et al.*, 2021).

Entretanto, a pesquisa em co-design apresenta seus próprios desafios. Problemas como falhas de comunicação, fadiga de participação e considerações éticas frequentemente surgem devido à diversidade de origens e níveis de especialização entre as partes interessadas. Esses desafios podem complicar o processo colaborativo, exigindo uma interação cuidadosa e o compromisso de manter um diálogo aberto. Além disso, a necessidade de governança e alocação de recursos eficazes continua sendo fundamental para o sucesso das iniciativas de co-design, influenciando sua capacidade de adaptação e resposta às contribuições das partes interessadas.

Sendra (2024) discute a crescente adoção do co-design em processos de regeneração urbana, destacando a falta de definições claras em políticas públicas e propondo dez princípios éticos fundamentais para orientar os processos de codesign, garantindo que sejam conduzidos de forma justa e inclusiva. Primeiramente, a autonomia é enfatizada, garantindo que as partes interessadas tenham a liberdade de tomar suas próprias decisões sem coerção. A responsabilidade é outro princípio essencial, enfatizando que todos os participantes devem agir com integridade e considerar as consequências de suas ações. A transparência também demonstra-se importante, exigindo que o processo seja claro, com informações acessíveis

SUMÁRIO

sobre as intenções e limitações do projeto. Além disso, o respeito à diversidade, promovendo a inclusão de diferentes culturas, identidades e perspectivas, enquanto a participação genuína garante que todos os contribuintes possam se envolver significativamente no desenvolvimento do projeto. O princípio do empoderamento visa permitir que as comunidades influenciem os resultados, garantindo um processo verdadeiramente transformador. A justiça é igualmente relevante, com foco na equidade, garantindo acesso igualitário a oportunidades e recursos. A confiança mútua deve ser cultivada, criando um ambiente seguro para a troca de ideias, enquanto a sustentabilidade é enquadrada como um compromisso com soluções de longo prazo que considerem os impactos sociais, ambientais e econômicos. Por fim, a reflexividade incentiva os participantes a avaliarem continuamente suas próprias práticas e decisões, ajustando-as conforme necessário para garantir que o processo de co-design permaneça ético e responsável ao longo de sua execução. Quando seguidos, esses princípios garantem que o co-design seja eficaz, justo e eticamente sólido.

Nos últimos anos, a aplicação de metodologias de co-design resultou em histórias de sucesso significativas, demonstrando sua eficácia no desenvolvimento de soluções centradas no usuário e no fomento do engajamento da comunidade. À medida que o co-design continua a evoluir, pesquisas em andamento buscam abordar seus desafios inerentes, promovendo práticas sustentáveis e inclusivas que garantam que todas as vozes sejam ouvidas e valorizadas no processo de design. A publicação “The Future of Competition: Co-Creating Unique Value with Customers” é frequentemente creditada por popularizar os conceitos de co-criação e co-design dentro das organizações, incentivando-as a envolver ativamente os usuários no processo de design, em vez de apenas buscar sua contribuição de forma consultiva (Prahalad & Ramaswamy, 2004). Desde então, o movimento amadureceu e se tornou uma estrutura de design colaborativo, cujo objetivo é desenvolver soluções inovadoras integrando as ideias e experiências de diversos membros da comunidade desde o início do processo de design (Benz *et al.*, 2024).

SUMÁRIO

Por meio desse engajamento colaborativo, as partes interessadas deixam de ser meros receptores passivos dos resultados do design; tornam-se participantes ativos em todas as etapas, desde a identificação do problema até a ideiação, prototipagem e avaliação. Essa abordagem visa aprofundar a empatia e a compreensão das necessidades do usuário, permitindo a criação de soluções que sejam não apenas relevantes, mas também práticas e inovadoras (Busciantella-Ricci & Scataglini, 2024). À medida que as práticas de co-design ganham impulso, elas moldam as metodologias empregadas em diversos setores, garantindo que os processos de design sejam baseados em contextos do mundo real e nas aspirações da comunidade.

Embora os princípios específicos do codesign possam variar (Southwick *et al.*, 2020; Javanparast *et al.*, 2022), vários temas-chave emergem consistentemente:

Inclusão: O envolvimento de todas as partes interessadas, como designers, usuários finais e membros da comunidade, garante que diversas perspectivas sejam incorporadas ao processo de design.

Empoderamento: O co-design permite que os indivíduos contribuam significativamente, envolvendo diretamente os usuários, permitindo que as suas ideias moldem os resultados finais.

Desenvolvimento Iterativo: O co-design enfatiza uma abordagem cíclica na qual as ideias são continuamente refinadas por meio de feedback e testes. Esse processo iterativo é crucial para adaptar as soluções às necessidades do usuário.

Propriedade Compartilhada: O co-design promove um senso de responsabilidade compartilhada entre os participantes, pois todos têm interesse nos resultados do design.

SUMÁRIO

O co-design emprega uma variedade de técnicas e ferramentas para facilitar a colaboração entre as partes interessadas e gerar soluções eficazes para os problemas identificados. Esses métodos variam de workshops à prototipagem e diversas estratégias de facilitação projetadas para aumentar a participação e o engajamento.

Workshops colaborativos servem como uma técnica primária no co-design, permitindo que as partes interessadas se envolvam ativamente em sessões interativas de resolução de problemas. Esses workshops geralmente duram várias horas e são estruturados para incentivar o pensamento criativo e a geração de ideias (Lee, Roberts & Taormina, 2013; Wong *et al.*, 2021). A natureza interativa dessas sessões aumenta o engajamento em comparação com entrevistas ou pesquisas tradicionais, embora possam exigir muitos recursos e um comprometimento significativo de tempo dos participantes.

A prototipagem é outra ferramenta essencial no processo de co-design, frequentemente utilizada para transformar ideias em representações tangíveis que podem ser testadas e refinadas iterativamente (Walker *et al.*, 2023). Diversas formas de prototipagem são empregadas, incluindo protótipos de baixa fidelidade, artefatos visuais e mock-ups digitais. Essa abordagem facilita o teste de conceitos e aprimora a compreensão dos participantes sobre as soluções em desenvolvimento.

Além disso, com a crescente tendência de colaboração remota, formatos híbridos que combinam sessões presenciais e online tornaram-se cada vez mais populares no co-design. Essa abordagem permite maior flexibilidade e inclusão, acomodando as diferentes preferências e disponibilidades dos participantes. Ferramentas digitais também podem facilitar a colaboração, fornecendo plataformas para brainstorming, feedback e desenvolvimento iterativo de ideias. A combinação de ferramentas físicas e digitais permite uma estratégia de engajamento mais abrangente, adaptando-se às necessidades dos participantes e aumentando a eficácia metodológica.

SUMÁRIO

Apresentamos três estudos de caso que incorporam o co-design: o primeiro, de Priscilla Van Even, “Co-design de uma Ferramenta Interativa e Prática para Contar Histórias Científicas”. O segundo, de Jonas Ferrigolo Melo, “Desenvolvimento de um Modelo de Gerenciador de Consentimento Ético: Uma Abordagem Participativa para a Proteção de Dados Pessoais em Websites”. E o terceiro, de Luciana Monteiro-Krebs, intitulado “Co-design de interface para mediação algorítmica de notícias”. Esses casos ilustram como os princípios do co-design podem ser aplicados de forma eficaz em diferentes domínios, demonstrando seu potencial para promover soluções mais inclusivas, éticas e centradas no usuário.

CO-DESIGN UMA FERRAMENTA INTERATIVA E PRÁTICA PARA CONTAR HISTÓRIAS CIENTÍFICAS

A crescente midiatização e as tecnologias emergentes mudaram a forma como os cientistas comunicam suas pesquisas. Isso cria oportunidades para novas práticas criativas e interativas de comunicação científica e para a popularização da ciência, mas também apresenta riscos como a diminuição da qualidade da comunicação e a propagação de desinformação (Van Even, 2023). No âmbito do projeto ParCos (Participatory Communication of Science – Financiado pelo Horizonte 2020), buscamos explorar e expandir o escopo da narrativa científica participativa, ao mesmo tempo em que aprimoramos as práticas de narrativa científica e garantimos a confiabilidade e a qualidade da informação científica compartilhada.

Para atingir esses objetivos, desenvolvemos uma estrutura de avaliação com critérios de qualidade robustos e baseados em evidências para avaliar a comunicação científica criativa e interativa

SUMÁRIO

para fins de divulgação pública (Van Even, Zaman e Hannes, 2021). Essa estrutura de avaliação consolida os resultados de uma revisão sistemática da literatura, mapeando e identificando os critérios de qualidade existentes para apoiar os comunicadores científicos e aprimorar a narrativa científica. A revisão sistemática e a estrutura de avaliação com critérios detalhados podem ser consultadas em Van Even *et al.* (2025, no prelo). O protocolo que sustenta a revisão está disponível em Van Even, Zaman e Hannes (2020).

A estrutura de avaliação foi desenvolvida para apoiar os comunicadores científicos na avaliação de seus próprios esforços de divulgação científica, bem como dos seus pares, com o objetivo de evitar ou mitigar o risco de comunicação inadequada, não confiável ou malsucedida em informações científicas, mediando e fomentando, assim, a confiança na ciência. Além disso, buscamos apoiar os comunicadores científicos no desenvolvimento proativo de uma divulgação científica interativa e criativa.

No entanto, embora a estrutura tenha se mostrado útil para avaliar os esforços de comunicação científica pública, identificamos uma lacuna prática: os comunicadores científicos não foram incentivados a se envolver ativamente com ela. Essa constatação nos levou a desenvolver uma ferramenta amigável, envolvente e reflexiva para a concepção e avaliação de reportagens científicas, a fim de promover a adoção ativa dos critérios de qualidade. Empregamos uma abordagem de co-design, garantindo que o desenvolvimento da ferramenta envolvesse diretamente o usuário final, permitindo-nos traduzir a estrutura com os critérios em um formato mais criativo e interativo. Qual melhor maneira de desenvolver uma ferramenta prática do que incluir aqueles que a utilizariam no processo de desenvolvimento?

METODOLOGIA E APLICAÇÃO

ETAPA 1: DA ESTRUTURA DE AVALIAÇÃO AO PRIMEIRO PROTÓTIPO COM VALIDADE APARENTE

Nossa ferramenta foi co-projetada em três etapas consecutivas. Durante a primeira etapa, organizamos verificações de validade aparente com especialistas, que serviram tanto como uma avaliação qualitativa inicial da praticidade e relevância da estrutura quanto como um meio de gerar ideias concretas para sua tradução em uma ferramenta prática. A **validade aparente** é uma forma de avaliação qualitativa na qual especialistas e/ou usuários em potencial determinam se um instrumento ou ferramenta atinge seu propósito pretendido em contextos e aplicações do mundo real.

Essas verificações de validade aparente se concretizaram em **workshops de co-criação** altamente interativos, com duração de duas horas, com profissionais de comunicação científica e acadêmicos, o público-alvo da ferramenta. Esses workshops, com o objetivo de coletar contribuições com base na expertise de diferentes partes interessadas, ocorreram tanto online quanto presencialmente.

Organizamos cinco sessões online com pesquisadores e diversos profissionais, incluindo especialistas em interação humano-computador, profissionais de mídia, educadores de arte e profissionais de museus, representando um amplo espectro de expertise. Os participantes eram de várias localizações geográficas, como Bélgica, Egito, Dinamarca, Finlândia, Malta, Suécia e Reino Unido. Essas oficinas foram facilitadas no Miro. O **Miro** é uma plataforma de quadro branco digital para uma variedade de tarefas que exigem visualização, criatividade, organização e colaboração. Além disso, organizamos duas sessões interativas presenciais na Bélgica com alunos

SUMÁRIO

de bacharelado e mestrado. Uma sessão ocorreu na KU Leuven e a outra na Universidade de Antuérpia. Esses alunos participaram de maneira semelhante às sessões online, envolvendo-se em atividades reflexivas e discussões. A inclusão de alunos garantiu que a ferramenta fosse acessível e relevante também para futuros profissionais.

O conceito básico das oficinas interativas foi envolver todos os participantes na reflexão sobre a estrutura de avaliação através das lentes de suas próprias práticas de narrativa científica. Os participantes foram convidados a compartilhar seus conhecimentos e experiências com a estrutura e a propor sugestões específicas para torná-la mais prática e fácil de usar. Essa abordagem nos permitiu avaliar qualitativamente suas percepções sobre a estrutura e seus critérios de qualidade, identificar quaisquer lacunas percebidas e explorar suas ideias para uma ferramenta envolvente e acessível.

Durante as sessões, os participantes anotaram seus pensamentos e sugestões individuais em post-its (virtuais ou físicos) em resposta a perguntas específicas. Essa contribuição individual foi seguida por discussões em grupo, nas quais os participantes trocaram perspectivas em voz alta. Essas discussões permitiram a geração colaborativa de ideias, visualizações e refinamentos adicionais para traduzir a estrutura em uma ferramenta prática.

Essa abordagem provou ser uma excelente maneira de facilitar efetivamente o intercâmbio de diversas formas e tipos de expertise. A natureza iterativa dessas oficinas incentivou um envolvimento mais profundo com o material, resultando em insights mais ricos. As contribuições foram coletadas, analisadas e sintetizadas pelo pesquisador principal. Esse processo ajudou a traduzir a estrutura de avaliação, inicialmente rígida e teórica, com critérios de qualidade, em um primeiro **protótipo** utilizável: um conjunto de cartões com diretrizes práticas, projetado para apoiar comunicadores científicos no desenvolvimento de histórias científicas eficazes para o público em geral.

SUMÁRIO

O processo de validade aparente e os resultados foram documentados em Van Even (2023, p. 339) e em Van Even, Zaman e Hannes (2022a).

Etapa 2: Iterações e ajustes do protótipo

Em uma segunda etapa, operacionalizamos o protótipo e organizamos quatro **sessões interativas** de experimentação com alunos do ensino médio, desde o bacharelado até o doutorado. Três dessas sessões ocorreram na KU Leuven, na Bélgica, e uma foi realizada durante um curso de verão internacional na VIU, em Veneza. As sessões de experimentação foram organizadas como workshops de quatro horas, nos quais os alunos trabalharam em grupos de quatro a cinco participantes para criar e iterar histórias científicas usando o protótipo de cartas.

As cartas continham perguntas norteadoras na parte inferior, elaboradas para estimular reflexões individuais e/ou discussões em grupo. Essas perguntas eram diretamente influenciadas pelos critérios de qualidade da estrutura de avaliação. As cartas focavam em quatro dimensões básicas da narrativa científica: (1) o contexto científico da história, (2) a tradução e a apresentação da história, (3) o uso de uma ferramenta tecnológica para apresentar a história e (4) o impacto da história. Cada dimensão foi apresentada com diferentes níveis de dificuldade, acomodando participantes com diferentes níveis de conhecimento, do iniciante ao avançado.

Durante esses testes, os participantes não apenas criaram em conjunto histórias científicas abrangentes, sólidas e eficazes usando as cartas, mas também testaram o protótipo e forneceram feedback sobre seu conteúdo e formato. Os testes nos permitiram identificar os pontos fortes da ferramenta, apontar áreas para melhorias e revelar potenciais barreiras ou ambiguidades em seu uso bem-sucedido. As experiências e reflexões compartilhadas pelos alunos foram essenciais para o aprimoramento do protótipo. Essas ideias

SUMÁRIO

orientaram uma nova iteração da ferramenta, garantindo que ela fosse mais intuitiva, envolvente e adaptável a diversos cenários de comunicação científica.

O processo foi documentado em Van Even (2023, pp. 340-341), enquanto as iterações e resultados foram relatados em Van Even, Zaman e Hannes (2022b, c, d).

Etapa 3: Finalização do protótipo, avaliação da ferramenta interativa

Na terceira e última fase, os achados das verificações de validade aparente e das sessões de teste foram integrados às anotações de campo obtidas por meio de **observação participativa**, um método antropológico. Essa abordagem observacional proporcionou percepções mais aprofundadas, contextualizando os comentários e sugestões feitos pelos participantes nas etapas anteriores. Com base nessas observações de campo, o protótipo foi ainda mais aprimorado e refinado, resultando na versão final de uma ferramenta prática chamada 'Trainer Cards' (ver Figura 1) e em seu manual complementar.

Figura 1 - A ferramenta 'Trainer Cards' para Narrativa Científica



Como parte dessa fase de finalização, a ferramenta interativa passou por mais uma rodada de testes de validade aparente, conduzida pelo consórcio do projeto ParCos. Esse consórcio incluiu acadêmicos e profissionais de mídia que aplicaram a ferramenta

SUMÁRIO

para avaliar suas histórias científicas. A aplicação prática e o feedback foram fundamentais para validar a praticidade e a eficácia da ferramenta final. O processo e os resultados desta validação foram documentados em Van Even, Zaman e Hannes (2022c). A versão final dos Cartões de Treinamento e seu manual podem ser encontrados em Van Even (2024).

Em conclusão, podemos afirmar que, ao final desta jornada de co-design, foi co-desenhada uma ferramenta interativa para apoiar comunicadores científicos na elaboração, avaliação e aprimoramento de suas narrativas científicas voltadas para públicos mais amplos. Para além de sua utilidade prática, essa ferramenta também incorpora a filosofia do co-design, ao incentivar práticas colaborativas na criação de histórias sobre ciência. As Trainer Cards são, portanto, não apenas um produto do co-design, mas também uma ferramenta que promove ativamente metodologias de co-design na comunicação científica. As Trainer Cards são, assim, uma ferramenta 'feita por' e 'para' o co-design.

LIÇÕES APRENDIDAS

Em algumas ocasiões, os participantes podem dizer aquilo que acreditam que você ou outras pessoas gostariam de ouvir, o que pode influenciar o processo de design e impactar os resultados. Por isso, é importante pedir esclarecimentos aos participantes, explorar seus motivos e promover ativamente um ambiente em que se sintam à vontade para expressar opiniões sinceras.

Ao organizar um processo de co-design, participantes de diferentes origens não apenas contribuem com seus conhecimentos específicos, mas também adquirem percepções importantes uns com os outros. Eles entram em contato com novas perspectivas e

SUMÁRIO

aspectos que talvez nunca tivessem considerado. Esse aprendizado recíproco torna o processo de co-design benéfico não apenas para o pesquisador responsável ou para o desenvolvimento da ferramenta, mas também para os próprios participantes.

Na narrativa científica, é particularmente importante integrar várias dimensões: (1) a base científica da história, (2) a comunicação e apresentação das ideias, (3) o uso de ferramentas tecnológicas e (4) o impacto da história no público. Essas dimensões muitas vezes são tratadas separadamente na prática, dependendo da expertise do comunicador científico. No entanto, ao co-desenhar os cartões, esses elementos são reunidos em uma estrutura unificada, promovendo uma abordagem mais holística para a narrativa científica.

Coordenar o co-design é complexo e pode ser altamente exigente. Uma coordenação inadequada pode levar ao desengajamento de alguns participantes ou resultar em resultados ineficientes. Para lidar com isso, uma preparação cuidadosa é fundamental. Os facilitadores devem antecipar possíveis barreiras de comunicação, estabelecer objetivos claros e observar ativamente a dinâmica durante o processo para garantir oportunidades de participação igualitária. Ao trabalhar em grupos, alguns participantes podem naturalmente assumir papéis mais dominantes. Como moderador, é essencial equilibrar essas dinâmicas, incentivando os participantes mais silenciosos a contribuir e garantindo que todos tenham voz. Uma facilitação inadequada pode levar à falta de diversidade nos resultados e a uma representação incompleta das perspectivas.

SUMÁRIO

DESENVOLVIMENTO DE UM MODELO ÉTICO DE GERENCIADOR DE CONSENTIMENTO: UMA ABORDAGEM PARTICIPATIVA PARA A PROTEÇÃO DE DADOS PESSOAIS EM WEBSITES

Em resposta às crescentes denúncias de uso indevido de informações pessoais, a opinião da sociedade sobre proteção de dados evoluiu e ganhou nuances nas últimas décadas (PARKER; PINE; ERNST, 2018), seja pelo aumento das legislações de proteção de dados, pela intensificação das invasões de privacidade ou até mesmo por decisões judiciais sobre vazamentos de dados. Apesar da ampliação dos problemas relacionados à privacidade, os Estados Unidos optaram por tratar essas questões por meio de esquemas regulatórios fragmentados que protegem apenas certos tipos de informação, em vez de legislações abrangentes como o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia (PARKER; PINE; ERNST, 2018) e a Lei Geral de Proteção de Dados Pessoais (LGPD), no Brasil (BRASIL, 2018).

Na Europa, os órgãos reguladores da privacidade foram mais bem-sucedidos ao instituírem restrições à coleta de dados desde o início. Em 1995, a Diretiva de Proteção de Dados da União Europeia estabeleceu exigências precisas para um consentimento “inequívoco” antes do tratamento de informações pessoais e determinou que os dados deveriam ser coletados para “fins específicos, explícitos e legítimos” (COMISSÃO EUROPEIA, 1995). Em 2002, essa diretiva foi sucedida pela Diretiva de Privacidade Eletrônica – *ePrivacy* (PARLAMENTO EUROPEU, 2002), que passou a exigir informações “claras e abrangentes” antes do armazenamento de cookies (CHANDER, 2013). Os avisos de consentimento para cookies surgiram com a Diretiva 2009/136/EC da União Europeia (PARLAMENTO

SUMÁRIO

EUROPEU, 2009), que alterou o artigo 5 da *ePrivacy* de 2002, exigindo que dados fossem armazenados nos dispositivos dos controladores apenas após o consentimento do usuário com base em “[...] informações claras e abrangentes [...] sobre as finalidades do tratamento” (PARLAMENTO EUROPEU, 2009). Mais recentemente, uma diretiva da UE de 2011 concedeu aos usuários o direito de recusar o uso de cookies (COMISSÃO EUROPEIA, [s.d.]).

Em 25 de maio de 2018, entrou em vigor o GDPR, representado pelo Regulamento (UE) 2016/679. O artigo 6 do regulamento define as bases legais para o tratamento de dados pessoais de residentes europeus, incluindo que “o titular dos dados deu o seu consentimento para o tratamento dos seus dados pessoais para uma ou mais finalidades específicas” (PARLAMENTO EUROPEU, 2016, artigo 6.1a). O Considerando 32 do GDPR e as diretrizes publicadas pelas autoridades europeias de proteção de dados (ARTICLE 29 DATA PROTECTION WORKING PARTY, 2011) exigem que o consentimento válido seja um ato afirmativo claro “que indique uma manifestação livre, específica, informada e inequívoca da vontade do titular dos dados, pela qual este, por meio de declaração ou ação afirmativa clara, concorda com o tratamento dos dados pessoais que lhe dizem respeito” (PARLAMENTO EUROPEU, 2016).

A LGPD, sancionada em 14 de agosto de 2018 e em vigor desde 18 de setembro de 2020, é um marco importante no ordenamento jurídico brasileiro ao regulamentar o tratamento de dados pessoais no país. Inspirada no GDPR, a LGPD estabelece princípios, direitos e deveres para empresas e autoridades públicas, buscando garantir maior transparência e controle aos titulares sobre o uso de suas informações. Num desdobramento histórico posterior, o Congresso Nacional promulgou a Emenda Constitucional 115/2022, que incluiu expressamente a proteção de dados pessoais no artigo 5º da Constituição Federal, elevando-a ao status de direito e garantia fundamental. Essa emenda reforça a proteção de dados como um direito essencial no Brasil, sublinhando sua importância para a cidadania em um contexto de crescente digitalização e uso intensivo de dados pessoais.

SUMÁRIO

De forma geral, as leis de proteção de dados buscam reduzir a assimetria de informação ao exigir, entre outras medidas, que as empresas informem aos titulares as finalidades para as quais seus dados pessoais estão sendo coletados. Para cumprir essa exigência, os controladores e operadores recorrem a políticas de privacidade, que servem como instrumentos formais de transparência. No entanto, os titulares ainda carecem de garantias suficientes quanto ao que, de fato, está sendo coletado sobre eles.

Isso evidencia como o problema da assimetria informacional é difícil de resolver, sobretudo devido aos custos impostos ao indivíduo: tempo, confiança, compreensão textual, interesse pela privacidade no ambiente digital, entre outros. As políticas de privacidade são frequentemente complexas demais e, na prática, a legislação de proteção de dados não resolve totalmente essa assimetria.

Jornalistas e pesquisadores frequentemente criticam a forma como as empresas de tecnologia tratam o consentimento. Quando buscam o consentimento para tratar dados pessoais, muitas empresas o fazem por meio de contratos, sendo o mais comum o EULA – Acordo de Licença de Usuário Final (SADOWSKI, 2019). Thatcher *et al.* (2016) definem esse tipo de contrato como uma marca registrada da tecnologia digital, presente na maioria das interações online – como aquelas páginas que obrigam o usuário a clicar em “Concordo” ou “Aceito” para continuar utilizando um serviço (SADOWSKI, 2019, p. 07).

Por serem aplicados genericamente a todos os usuários, os EULAs são considerados contratos de adesão (ZAMIR, 2014), redigidos em linguagem densa e jurídica (SADOWSKI, 2019). Ou seja, são menos um instrumento de consentimento real e mais um mecanismo de conformidade legal. Embora possam assumir outras formas além daquelas relacionadas diretamente à privacidade – como os Termos de Uso – neste trabalho, esses contratos são tratados sob a denominação genérica de “Política de Privacidade”.

SUMÁRIO

Parte-se da suposição de que os controladores de dados oferecem transparência por meio dessas políticas, termos e condições (CUSTERS *et al.*, 2013). O problema é que, na prática, as pessoas não leem essas políticas de privacidade (BORGESIUS, 2015; GROOM; CALO, 2011; OBAR; OELDORF-HIRSCH, 2020).

Mesmo aqueles que tentam lê-las enfrentam dificuldades com a linguagem jurídica e tecnológica. Em alguns casos, há políticas de privacidade mais claras e curtas, mas levanta-se a dúvida sobre a quantidade de informação efetivamente transmitida (TOUBIANA; NISSENBAUM, 2011). Para ilustrar o quanto é impraticável a leitura dessas políticas, um estudo estimou que, se os americanos lessem todas as políticas de privacidade palavra por palavra, o tempo gasto equivaleria a cerca de US\$ 781 bilhões por ano (MCDONALD; CRANOR, 2008). Na União Europeia, Oers e Groenendijk (2010) mostram que apenas 11% das pessoas leem o texto completo. McDonald e Cranor (2008) apontam que ler todas as políticas demandaria cerca de 244 horas por ano por pessoa, enquanto estudos indicam que as pessoas estão dispostas a gastar, em média, entre 1 e 5 minutos com esse tipo de leitura (VAN DEN BERG; VAN DER HOF, 2012).

Outro estudo concluiu que seriam necessários 76 dias úteis, trabalhando 8 horas por dia, para ler todas as políticas que uma pessoa encontra ao longo de um ano. O website *Visual Capitalist* analisou os serviços online mais populares e calculou quanto tempo seria necessário para ler seus termos de uso (LEPAN; VISUAL CAPITALIST, 2020). Com base em uma média de leitura entre 200 e 250 palavras por minuto, constataram que a maioria dos termos exigiria cerca de 30 minutos de leitura. Como referência, a Constituição dos Estados Unidos levaria 18 minutos para ser lida, enquanto os Termos de Serviço da *Microsoft* ultrapassam 1 hora, o mesmo tempo necessário para ler *Macbeth*, de Shakespeare.

Diante desse cenário, está em fase de conclusão um projeto de pesquisa com o objetivo de apresentar um modelo inovador para um gerenciador de consentimento inequívoco para websites. Para isso,

SUMÁRIO

uma das etapas metodológicas adotadas foi a abordagem de co-design, que buscou captar as percepções dos usuários sobre os pontos fortes e fracos dos gerenciadores de consentimento mais comuns (UTZ *et al.*, 2019). Essa abordagem visou reunir *insights* sob a ótica dos usuários sobre as características de um gerenciador que proporcione maior controle e compreensão sobre o uso de seus dados, alinhando-se às demandas legais e éticas do contexto digital atual.

METODOLOGIA E APLICAÇÃO

As oficinas de co-design integraram uma abordagem participativa para envolver os participantes na reflexão e construção de propostas para gerenciadores de consentimento éticos em websites. Essa abordagem está alinhada com o design participativo, conforme destacado por Muller *et al.* (1991), que compreende efetivamente as necessidades dos usuários e aborda tensões de privacidade que podem passar despercebidas em análises tradicionais. Essa metodologia permite que os participantes influenciem as escolhas técnicas com base em seus valores e necessidades, promovendo maior aceitação e alinhamento ético das soluções desenvolvidas.

As oficinas tinham como objetivo principal explorar, de maneira interativa, como os participantes vivenciam e idealizam interações comportamentais com gerenciadores de consentimento. Utilizando técnicas de design participativo, como *brainstorming* e prototipagem em papel, as oficinas facilitaram a construção coletiva de *insights* sobre características desejáveis para gerenciadores de consentimento éticos.

Cada sessão foi estruturada em dois momentos centrais: exposição teórica e atividades práticas. No primeiro momento, os participantes foram introduzidos a tópicos como dados pessoais,

SUMÁRIO

economia da informação, capitalismo de vigilância, colonialismo de dados e ética intercultural da informação. A exposição teórica também incluiu resultados de pesquisas relevantes e documentos técnicos. Esse contexto teórico visava fundamentar as discussões e apoiar a fase prática.

Os participantes foram organizados em pequenos grupos e convidados a analisar 11 gerenciadores de consentimento durante o momento participativo. Eles identificaram pontos fortes e fracos usando etiquetas vermelhas (para fraquezas) e verdes (para pontos fortes), justificando suas avaliações com *post-it*. Posteriormente, os grupos discutiram e propuseram características para um gerenciador de consentimento ideal, alinhado aos princípios éticos apresentados anteriormente. Ao final, os grupos compartilharam suas conclusões em uma discussão coletiva.

Figura 2 - Imagens das oficinas



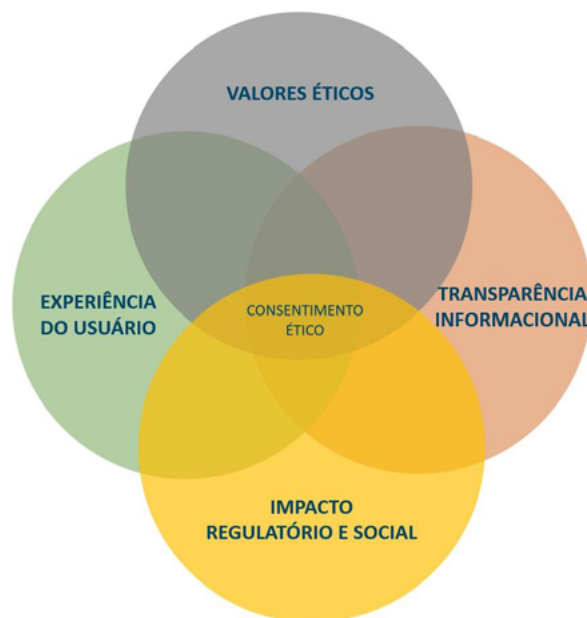
Entre março e outubro de 2024, foram realizadas cinco oficinas de co-design – uma em Portugal e quatro no Brasil – com a participação de 70 pessoas. A coleta de dados resultante dessas oficinas envolveu a organização das perspectivas expressas pelos participantes. Gerenciadores de consentimento com características semelhantes foram agrupados de acordo com suas avaliações,

SUMÁRIO

o que permitiu identificar convergências e divergências nas opiniões, bem como estabelecer padrões nas respostas. Esse processo reflete a visão de design orientado à pesquisa proposta por Fallman (2003), que busca integrar prática e investigação, gerando contribuições holísticas capazes de equilibrar perspectivas potencialmente conflitantes. Ao adotar uma abordagem participativa, as oficinas demonstraram a relevância do envolvimento direto dos usuários na construção de soluções que combinem inovação técnica com valores éticos e sociais.

O método das oficinas de co-design adotado neste projeto mostrou-se essencial para alcançar o objetivo geral da pesquisa: desenvolver um modelo de gerenciador de consentimento ético. A experiência prática e interativa das oficinas gerou *insights* que permitiram uma análise detalhada das expectativas, percepções e necessidades dos participantes em relação a esse tipo de ferramenta. Esses resultados compõem o que chamamos de "*tool-to-think-with*", conceito inspirado na abordagem de Frauenberger *et al.* (2015), que integra práticas reflexivas e críticas ao processo de design participativo. Esta ferramenta orientadora fornece elementos reflexivos que auxiliam tanto no design de soluções mais éticas quanto nas decisões metodológicas, analíticas e projetuais da pesquisa. Os materiais gerados pelas oficinas revelaram pressupostos implícitos e questões muitas vezes não verbalizadas, influenciando diretamente o desenvolvimento do modelo proposto, especialmente no que se refere à transparência da linguagem, à estética e às práticas que reforçam ou mitigam desigualdades informacionais. Estruturada em torno de quatro dimensões – valores éticos, experiência do usuário, transparência informacional e impacto regulatório e social (Figura 2) – a "*tool-to-think-with*" permitiu analisar criticamente os dados e garantir que o modelo de consentimento proposto incorporasse as qualidades desejadas pelos participantes, promovendo práticas mais justas e respeitadas à privacidade em ambientes digitais.

Figura 3 - Dimensões de “tool-to-think-with”



A escolha das quatro lentes propostas para a estruturação do modelo de gerenciador de consentimento se baseia nos resultados das oficinas de co-design; na necessidade de garantir um sistema que seja ao mesmo tempo ético, funcional e alinhado às exigências legais e sociais; e nas características do consentimento informado apresentados pela Lei Geral de Proteção de Dados Pessoais (LGPD), pelo Regulamento de Proteção de Dados Pessoais (GDPR), e pelos modelos conceituais de consentimento informado de Levine (1975), Meisel e Roth (1981), Faden, Beauchamp e King (1986), Ahmed *et al.* (2020) e Mayne, Howitt e Rennie (2016).

Com os dados coletados a partir destas oficinas será possível desenvolver um modelo de gerenciador que respeite, de fato, a opinião dos usuários sobre a coleta de seus dados pessoais, buscando oferecer uma abordagem inovadora para o consentimento, alinhada tanto às exigências legais quanto às necessidades reais dos usuários.

SUMÁRIO

Sendo assim, a “tool-to-think-with”, funcionará como um guia visual e conceitual, permitindo que pesquisadores, designers e desenvolvedores analisem criticamente a estrutura e os princípios que devem orientar um gerenciador de consentimento ético e eficaz.

LIÇÕES APRENDIDAS

As oficinas de co-design realizadas ao longo desta pesquisa forneceram *insights* e lições fundamentais para o desenvolvimento de um gerenciador de consentimento ético. Ao promover a colaboração entre participantes e pesquisadores, essas sessões revelaram tanto os pontos fortes quanto as fragilidades dos gerenciadores de consentimento existentes, além de evidenciar diferentes percepções dos usuários sobre privacidade e controle de dados. Os principais aprendizados destacaram a importância da clareza e da acessibilidade nas interfaces de consentimento, sendo o uso de jargões técnicos e linguagem excessivamente complexa apontado como barreira à compreensão. Os participantes também expressaram o desejo por maior controle granular sobre os dados compartilhados e por interfaces que reflitam intuitivamente suas preferências. Confiança e transparência surgiram como fatores decisivos para a disposição em consentir, com ênfase na necessidade de explicações claras sobre o uso e armazenamento dos dados. Além disso, foi destacada a importância de equilibrar estética e funcionalidade, evitando práticas manipulativas de design, tal como os *dark patterns*.

Apesar dos resultados positivos, o processo enfrentou desafios que exigiram estratégias adaptativas. A diversidade de experiências e formações dos participantes dificultou, em alguns momentos, a convergência de perspectivas, sendo necessário organizar os grupos de forma interdisciplinar. Barreiras linguísticas também surgiram, especialmente com termos técnicos e vocabulário específico de

SUMÁRIO

design, o que foi atenuado com o uso de recursos visuais e exemplos práticos pelos facilitadores. Restrições de tempo limitaram a profundidade das discussões, sendo priorizadas atividades-chave como *brainstorming*, prototipagem e momentos de reflexão coletiva.

Outro desafio foi o desengajamento de alguns participantes, que poderiam ter sido mais envolvidos com a oferta de incentivos, como compensação financeira. No geral, a abordagem de co-design se mostrou essencial para captar necessidades reais e expectativas autênticas dos usuários, garantindo que as soluções propostas fossem alinhadas aos valores éticos da pesquisa, como transparência, justiça e empoderamento. As oficinas também geraram aprendizados metodológicos que podem ser aplicados em outros contextos centrados no usuário e na ética de dados. O processo evidenciou, assim, o potencial do co-design para aproximar o design técnico das necessidades dos usuários, sobretudo em aplicações voltadas à privacidade.

CO-DESIGN COMO FERRAMENTA DE ELICITAÇÃO DE NECESSIDADES E PROTOTIPAGEM DE INTERFACES CENTRADAS NO USUÁRIO: A TÊNUE LINHA ENTRE TRANSPARÊNCIA E CONTROLE NA MEDIAÇÃO ALGORÍTMICA DE NOTÍCIAS NO FACEBOOK

Sistemas algorítmicos estão cada vez mais presentes em nosso cotidiano, influenciando o comportamento das pessoas e a dinâmica de governança em diversos domínios sociais (WILLSON, 2017). No contexto das notícias, a distribuição depende cada vez mais de sistemas de recomendação algorítmicos, projetados para prever as preferências dos leitores e oferecer uma experiência personalizada.

SUMÁRIO

No entanto, o uso crescente de sistemas de recomendação levanta preocupações sobre possíveis problemas, como vieses em processos de filtragem e personalização, falta de transparência, acesso limitado a pontos de vista plurais e diversos, e questões éticas relacionadas à opacidade, discriminação e desafios à autonomia dos usuários.

Para mitigar alguns desses problemas, legisladores em diferentes países e regiões têm buscado impor legalmente a transparência dos sistemas automatizados de tomada de decisão. Na Europa, o Regulamento Geral de Proteção de Dados (GDPR) estabelece obrigações de transparência para os controladores de dados (como plataformas de mídia social e provedores de notícias) e concede aos “titulares dos dados” (ou seja, usuários desses serviços) o direito de solicitar informações e controlar o uso de seus dados. Nesse contexto, as políticas de privacidade e os recursos de transparência tornam-se elementos centrais.

Pesquisas anteriores, no entanto, demonstraram que as políticas de privacidade são frequentemente ineficazes, incompreensíveis e ignoradas pelos usuários, além de desconectadas dos requisitos legais que justificam sua existência. Nesse sentido, este artigo investiga como os usuários percebem e se relacionam com o desafio da transparência em sistemas de recomendação, não como “titulares de dados”, mas como usuários. O objetivo é explorar como o *design* da interface pode aumentar a transparência desses sistemas, envolvendo os usuários no processo de *design*.

Pesquisas anteriores, no entanto, demonstraram que as políticas de privacidade são frequentemente ineficazes, incompreensíveis e ignoradas pelos usuários, além de desconectadas dos requisitos legais que justificam sua existência. Nesse sentido, este caso demonstra como os usuários de redes sociais percebem e se relacionam com o desafio da transparência em sistemas de recomendação, não como “titulares/proprietários de dados”, mas como usuários. O objetivo foi

SUMÁRIO

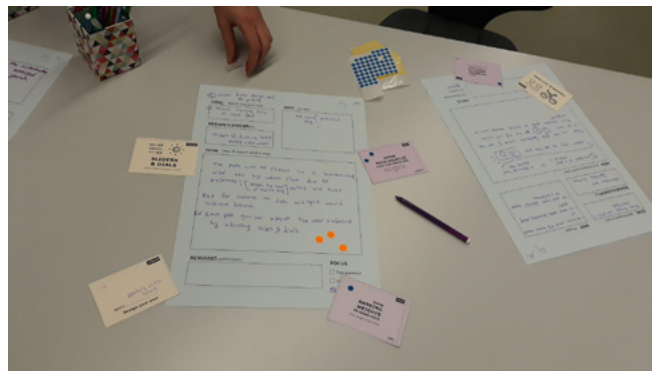
explorar como o *design* da interface pode aumentar a transparência desses sistemas, envolvendo os usuários no processo de *design*. O estudo foi conduzido no *Meaningful Interactions Lab* (Mintlab), na KU Leuven (Bélgica), em um esforço multidisciplinar parte do projeto de pesquisa *Algorithmic Transparency and Accountability in Practice* (ATAP), com mais detalhes por Monteiro-Krebs *et al.* (2019), Naudts *et al.* (2020) e Storms, Alvarado e Monteiro-Krebs (2022).

METODOLOGIA E APLICAÇÕES

A pesquisa foi conduzida em duas fases. Na primeira fase, foram organizados *workshops* de *co-design* para (1) investigar os “imaginários” dos usuários sobre os sistemas de recomendação algorítmicos e (2) convidá-los a idealizar elementos de interface para aumentar a transparência. A escolha por *workshops* de *co-design* se deu pela característica da metodologia de considerar os usuários como parceiros e envolvê-los como cocriadores no processo de *design*. Na segunda fase, algumas das ideias geradas nos *workshops* foram transformadas em um protótipo, levando em consideração os requisitos do GDPR, e foi realizada uma avaliação qualitativa desse protótipo com um novo grupo de participantes.

SUMÁRIO

Figura 4 - Materiais usados no workshop de co-design



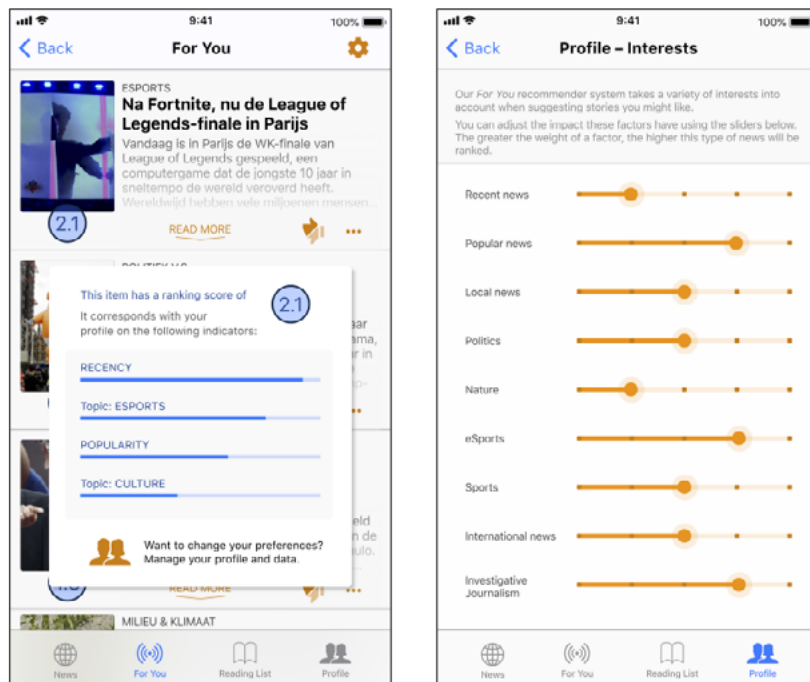
Fonte: os autores.

Para o primeiro *workshop*, foram recrutados usuários do Facebook, buscando abranger usuários de um serviço popular e que integra a classificação algorítmica em sua essência.¹ Os participantes realizaram uma atividade preparatória, na qual registraram as notícias que encontraram em seus *feeds* do Facebook, e participaram de um *brainstorming* coletivo para discutir suas experiências e percepções sobre como o Facebook seleciona e ordena as notícias.¹ No segundo *workshop*, os participantes realizaram exercícios de *co-design* para desenvolver recursos de explicabilidade que melhorassem a compreensão de como e por que certos itens aparecem em seus *feeds* de notícias.

Os principais resultados dos *workshops* de *co-design* incluem a identificação de diferentes tipos de informações que os usuários desejam ver sobre o funcionamento dos sistemas de recomendação, como os dados usados para personalizar as notícias, as razões pelas quais certos itens são recomendados e a possibilidade de influenciar as recomendações. Os participantes também sugeriram ideias para o *design* de interfaces, como painéis de controle que permitem aos usuários ajustar suas preferências e fornecer *feedback* sobre as recomendações (Figura 5).

SUMÁRIO

Figura 5 - Sobreposição e painel de controle para explicação baseada em pontuação

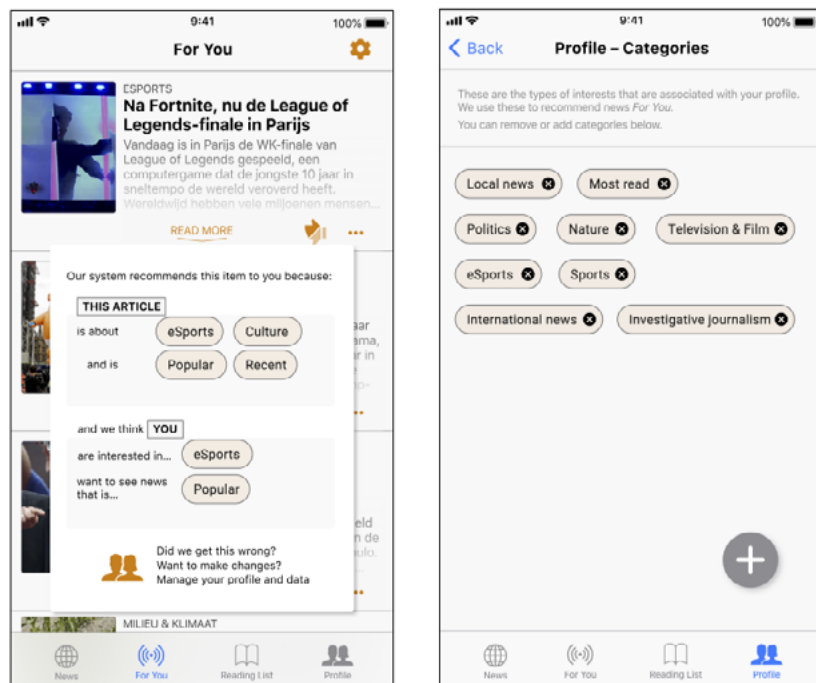


Fonte: (STORMS, ALVARADO, MONTEIRO-KREBS, 2022, p. 405:13)

O protótipo desenvolvido com base nas ideias dos *workshops* incluiu recursos como um painel de privacidade, um painel de controle com *sliders* para ajustar as preferências do usuário, botões de votação para fornecer *feedback* sobre as notícias e explicações sobre os fatores que influenciam as recomendações (Figura 6).

SUMÁRIO

Figura 6 - Sobreposição e painel para explicação baseada em “Por que isso?”



Fonte: Fig. 1 (STORMS, ALVARADO, MONTEIRO-KREBS, 2022, p. 405:12)

LIÇÕES DO CO-DESIGN PARA MEDIAÇÃO ALGORÍTMICA EM NOTÍCIAS

Nessa seção discutimos algumas das lições aprendidas ao praticar o co-design junto a usuários de plataformas digitais, tendo como objetivo principal aumentar a transparência e responsabilidade na mediação algorítmica de notícias.

SUMÁRIO

COMBINANDO TRANSPARÊNCIA
E CONTROLE

Os resultados do primeiro *workshop* mostraram as distinções que os participantes fazem em como os dados são usados na classificação de recomendadores algorítmicos em mídias sociais. Mesmo que os participantes traçassem as linhas entre essas categorias de forma diferente e demonstrassem diferentes níveis de preocupação com o uso de seus dados, é essencial notar que todos estavam interessados em ter mais controle sobre a coleta e o uso de tais dados. Essa decisão não se aplicava apenas ao uso de dados que consideravam ilegítimo, mas também aos dados ‘explicitamente fornecidos’: os participantes queriam estar mais envolvidos em como os provedores de notícias usam esses dados para fornecer recomendações. Isso ficou claro no segundo *workshop*, onde foram questionados sobre os recursos de transparência que gostariam de ver.

Transparência e Legibilidade. Quase todos os participantes queriam saber por que a interface exibia os itens em uma ordem específica. Eles queriam ver, por exemplo, quais fatores influenciam a curadoria automatizada. Os participantes queriam saber quais dados o sistema usava para construir seu perfil e como esses dados eram obtidos ou derivados. Nas palavras de um dos participantes, eles queriam “ver a lógica por trás disso, por que certos paralelos são traçados”. Durante as discussões posteriores no *workshop*, alguns participantes enfatizaram que essas informações devem ser transparentes, acessíveis e perspicazes. Como um participante colocou, adicionar “letrinhas miúdas” não é suficiente. A informação compartilhada também deve ser compreensível. Em outras palavras: transparência deve implicar legibilidade. Essa preocupação ressoa com questões levantadas na literatura acadêmica.

SUMÁRIO

Não há Transparência sem Controle. Os participantes consideraram transparência e controle como interligados. Esse ponto estava implicitamente presente em muitas propostas de *design* e ressurgiu na discussão de encerramento. No nível mais básico, muitos participantes sentiram que deveriam ser capazes de intervir se o sistema fizesse inferências incorretas sobre eles. Os participantes também expressaram a necessidade de controle refinado sobre o funcionamento do próprio sistema de recomendação para influenciar como ele chega às suas decisões. A transparência também deve envolver maior controle em um nível mais conceitual. Como observou um participante, “se for ‘apenas para informação e você não puder mudar nada, isso seria frustrante para mim”. Outro participante colocou de forma ainda mais direta: “[Transparência] significa controle”. Em outras palavras, os participantes desejavam transparência e *insights* que pudessem usar. Transparência não tem sentido sem meios adequados de controle. Os *designers* devem, portanto, considerar ambos os conceitos como dois lados da mesma moeda.

Preocupações com a Criação de Perfis. Durante os dois *workshops*, os participantes expressaram interesse em entender o perfil resultante que o sistema fazia deles. Uma das razões é que eles temiam particularmente ser colocados em uma categoria muito estreita, resultando no recebimento de notícias com base em uma concepção limitada de seus interesses. Isso estava relacionado à preocupação dos participantes com a diversidade nas notícias: eles queriam continuar descobrindo conteúdos novos e inesperados, especialmente se ‘o sistema’ os atribuísse a categorias específicas. Alguns participantes sugeriram que os *designers* poderiam alcançar a diversidade adicionando um botão ‘me surpreenda’ para mostrar variedade ou por meio de filtros elaborados que permitiriam aos usuários indicar preferências explicitamente. As preocupações com a criação de perfis, no entanto, eram mais amplas do que o medo de ser colocado em uma categoria restrita. Quase todos os participantes estavam ansiosos para saber o que o sistema pensava deles,

SUMÁRIO

ou seja, quais categorias o sistema os atribuía e como chegava a essas conclusões. Eles queriam saber quais dados o sistema usava para construir seu perfil e de onde esses dados vinham. Um sub-grupo de participantes imaginou um *dashboard* detalhado que permitiria aos usuários recuperar e manipular tais informações. Outros grupos propuseram variações de um botão 'por que isso?'. Em uma versão, estes incluíam os pesos de classificação dos itens no *feed* de notícias, detalhando como o item de notícia mostrado se assemelha ao perfil dos usuários.

LIÇÕES DA AVALIAÇÃO QUALITATIVA DO PROTÓTIPO

A avaliação do protótipo revelou uma série de *insights* sobre as preferências e preocupações dos usuários em relação aos recursos de transparência e controle em sistemas de recomendação de notícias.

Transparência é Multifacetada. Os participantes demonstraram interesse em diferentes tipos de informações sobre o funcionamento do sistema de recomendação. Alguns queriam entender os fatores que influenciam a classificação das notícias, enquanto outros estavam mais interessados em saber como seus dados são usados para personalizar as recomendações. Essa diversidade de interesses destaca a importância de fornecer uma variedade de recursos de transparência para atender às diferentes necessidades dos usuários.

Controle Aumenta a Confiança. Os participantes valorizaram a capacidade de controlar suas preferências e influenciar as recomendações. Os recursos de controle, como os *sliders* para ajustar as preferências e os botões de votação para fornecer *feedback*, foram bem recebidos pelos participantes.

SUMÁRIO

CONCLUSÕES

À medida que o co-design continua a ganhar força em diversos setores, futuras pesquisas devem se concentrar em aprimorar as práticas colaborativas envolvendo uma ampla gama de partes interessadas. Tendências emergentes indicam o uso crescente de ferramentas digitais de colaboração, que facilitam a participação remota e assíncrona, expandindo assim o acesso e promovendo a inclusão no processo de design.

O desenvolvimento de pesquisas como estas destaca a importância de estabelecer laboratórios que apoiem atividades de co-design, entre outros métodos colaborativos. O Mintlab na KU Leuven é um laboratório interdisciplinar de pesquisa, liderado pela professora Bieke Zaman, focado em explorar interações significativas entre pessoas e tecnologias digitais por meio de uma abordagem colaborativa e aberta. O laboratório integra ciência, sociedade, arte e design para investigar como as tecnologias digitais influenciam o comportamento humano e as dinâmicas sociais. Seus projetos incluem pesquisas sobre videogames, consumo mediado de esportes, habilidades digitais de jovens e ética de dados, sempre com forte ênfase em métodos participativos e co-design. Os três estudos de caso mencionados, de Jonas Ferrigolo, Priscilla Van Even e Luciana Monteiro Krebs, foram desenvolvidos total ou parcialmente dentro do Mintlab, refletindo a abordagem participativa do laboratório e a colaboração interdisciplinar.

Integrar métodos participativos como o co-design na educação dos estudantes e no treinamento de pesquisadores pode expandir a gama de abordagens utilizadas em pesquisa e inovação. Incorporar essas práticas nos currículos acadêmicos e no desenvolvimento profissional proporcionaria aos pesquisadores as ferramentas para se envolver de maneira mais eficaz com a sociedade. À medida que o interesse pela ciência cidadã cresce, integrar o

SUMÁRIO

co-design nas políticas universitárias poderia aumentar o impacto social da pesquisa e fortalecer a colaboração entre a academia e a comunidade em geral.

Ao mesmo tempo, a implementação desses métodos pode ser ainda mais apoiada por estruturas institucionais. Estabelecer laboratórios de pesquisa interdisciplinares poderia fornecer um ambiente estruturado, mas adaptável, onde o co-design é uma das várias abordagens utilizadas para facilitar o engajamento entre pesquisadores e sociedade. Esses laboratórios poderiam servir como espaços para experimentação metodológica, apoiando tanto sua aplicação quanto o desenvolvimento de outras práticas colaborativas de pesquisa.

Essa transição para plataformas digitais destaca a necessidade de pesquisas que avaliem a eficácia dessas ferramentas em melhorar o engajamento das partes interessadas e a qualidade geral dos resultados do co-design. A pesquisa futura sobre co-design deve priorizar a inclusão, garantindo que os designs atendam a uma gama mais ampla de habilidades e origens. Além disso, integrar a sustentabilidade nos processos de co-design está se tornando cada vez mais importante. Pesquisas devem explorar como incorporar considerações ambientais e sociais nas metodologias de co-design, permitindo que os projetos atendam não apenas às necessidades dos usuários, mas também ao seu impacto na comunidade e no meio ambiente.

A incorporação de dados e ideias dos usuários no processo de co-design é uma área em expansão a ser explorada. Análises avançadas podem desempenhar um papel crucial em informar e aprimorar as decisões de design, permitindo uma compreensão mais detalhada das necessidades e preferências dos usuários.

Assim, surge a necessidade de estudos focados no desenvolvimento de estruturas e laboratórios que promovam estudos de usuários e temas em ambientes onde as iterações de co-design são aplicadas, incluindo a coleta e análise desses dados e estratégias para implementar pesquisas com abordagens interativas.

SUMÁRIO

A pesquisa em co-design também deve adotar metodologias interdisciplinares, particularmente em áreas como Ciência da Informação, Comunicação e setores de inovação como tecnologia e Inteligência Artificial, com foco no estabelecimento de estruturas para colaboração interdisciplinar que priorizem as contribuições dos usuários e estejam melhor alinhadas com os contextos do mundo real nos quais são implementadas e os problemas a serem resolvidos.

Por fim, há uma necessidade crítica de explorar abordagens de capacitação para as partes interessadas além da academia, com o objetivo de aprimorar suas contribuições nos processos de co-design, investigando as necessidades de treinamento e desenvolvendo recursos educacionais que capacitem partes interessadas diversas a se envolverem de forma significativa no co-design, democratizando assim o processo de design e aumentando seu impacto na sociedade.

REFERÊNCIAS

ABRASSART, C.; GAUTHIER, P.; PROULX, S.; MARTEL, M. D. Le design social: une sociologie des associations par le design? Le cas de deux démarches de codesign dans des projets de rénovation des bibliothèques de la Ville de Montréal. **Lien social et Politiques**, n. 73, p. 117-138, 2015.

AHMED, J. *et al.* GDPR Compliant Consent Driven Data Protection in Online Social Networks: A Blockchain-Based Approach. IEEE, mar. 2020.

ARTICLE 29 DATA PROTECTION WORKING PARTY. Opinion 15/2011 on the definition of consent. 2011. Disponível em: https://web.archive.org/web/20220401032324/https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2011/wp187_en.pdf. Acesso em: 10 abr. 2022.

BENZ, C. *et al.* Community-based participatory-research through co-design: supporting collaboration from all sides of disability. **Research Involvement and Engagement**, v. 10, n. 1, p. 47, 2024.

SUMÁRIO

BIRD, M. *et al.* A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. **Research Involvement and Engagement**, v. 7, p. 1-12, 2021.

BORGESIU, F. Z. Informed consent: We can do better to defend privacy. **IEEE Security and Privacy**, v. 13, n. 2, p. 103-107, 2015.

BRASIL. **Lei n. 13.708, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). 2018.

BUSCIANTELLA-RICCI, D.; SCATAGLINI, S. Research through co-design. **Design Science**, v. 10, e3, 2024.

BURKETT, I. **An introduction to co-design**. Sydney: Knode, 2012.

CESTARI, V. R. F. *et al.* Codesign of a care-educational app for people with heart failure: Design, prototyping and co-implementation. **Texto & Contexto-Enfermagem**, v. 31, e20220163, 2023.

CHANDER, A. How law made Silicon Valley. **Emory LJ**, v. 63, p. 639, 2013.

CUSTERS, B. *et al.* Informed consent in social media use—the gap between user expectations and EU personal data protection law. **SCRIPTed**, v. 10, p. 435, 2013.

EUROPEAN COMMISSION. Directiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de outubro de 1995 relativa à protecção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. **Jornal Oficial das Comunidades Europeias**, 1995. Disponível em: https://web.archive.org/web/20211029134615/https://www.ipvc.pt/wp-content/uploads/2021/01/Directiva-n.%C2%BA-95_46_CE-do-Parlamento-Europeu-e-do-Conselho-de-24-de-outubro-de-1995.pdf. Acesso em: 20 abr. 2022.

EUROPEAN COMMISSION. The Europa Web Guide. Disponível em: <https://web.archive.org/web/20220407081032/https://wikis.ec.europa.eu/display/WEBGUIDE/04.+Cookies+and+similar+technologies>. Acesso em: 26 abr. 2022.

EUROPEAN PARLIAMENT. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector. **Official Journal of the European Communities**, 2002.

SUMÁRIO

EUROPEAN PARLIAMENT. Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC, Directive 2002/58/EC and Regulation (EC) No 2006/2004. **Official Journal of the European Union**, 2009.

EUROPEAN PARLIAMENT. Regulation (EU) 2016/679 of the European Parliament and of the Council - General Data Protection Regulation. **Regulation (EU)**, v. 679, 2016.

FADEN, R. R.; BEAUCHAMP, T. L. **A history and theory of informed consent**. [s.l.] Oxford University Press, 1986.

FALLMAN, D. Design-oriented human-computer interaction. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY: ACM, 2003. p. 5-12.

FRAUENBERGER, C. *et al.* In pursuit of rigour and accountability in participatory design. **International Journal of Human-Computer Studies**, v. 74, p. 93-106, 2015.

GROOM, V.; CALO, R. Reversing the privacy paradox: An experimental study. **TPRC**, 2011.

JAVANPARAST, S. *et al.* Embedding research codesign knowledge and practice: Learnings from researchers in a new research institute in Australia. **Research Involvement and Engagement**, v. 8, n. 1, p. 71, 2022.

KIRK, J. *et al.* Challenges in co-designing an intervention to increase mobility in older patients: a qualitative study. **Journal of Health Organization and Management**, v. 35, n. 9, p. 140-162, 2021.

LEPAN, N.; VISUAL CAPITALIST. Visualizing the Length of the Fine Print, for 14 Popular Apps - Visual Capitalist. Visual Capitalis, abr. 2020.

LEE ROBERTS, R.; TAORMINA, M. Collaborative co-design for library workshops. **Behavioral & Social Sciences Librarian**, v. 32, n. 1, p. 46-56, 2013.

LEVINE, R. J. The nature and definition of informed consent in various research settings. Preliminary paper prepared for the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Washington, DC: US Department of Health, Education, and Welfare, 1975.

MATIAS, N. Co-design in a historical context. **MIT Center for Civic Media**, 23 dez. 2011. Disponível em: <https://civic.mit.edu/2011/12/23/co-design-in-a-historical-context/>.

MAYNE, F.; HOWITT, C.; RENNIE, L. Meaningful informed consent with young children: Looking forward through an interactive narrative approach. **Early Child Development and Care**, v. 186, n. 5, p. 673-687, 2016.

SUMÁRIO

MCDONALD, A. M.; CRANOR, L. F. The cost of reading privacy policies. **I/S: A Journal of Law and Policy for the Information Society**, v. 4, 2008.

MEISEL, A.; ROTH, L. H. What we do and do not know about informed consent. **Jama**, v. 246, n. 21, p. 2473-2477, 1981.

MONTEIRO-KREBS, Luciana; ALVARADO RODRIGUEZ, Oscar Luis; DEWITTE, Pierre; AUSLOOS, Jef; GEERTS, David; NAUDTS, Laurens; VERBERT, Katrien. Tell Me What You Know. In: Extended Abstracts of the 2019 CHI Conference, 2019, Glasgow. **Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems** - CHI EA '19. New York: ACM Press, 2019. p. 1-6. Disponível em: <https://dl.acm.org/doi/10.1145/3290607.3312808>. Acesso em 18 nov. 2024.

MULLER, M. J. *et al.* Privacy, anonymity and interpersonal competition issues identified during participatory design of project management groupware. **ACM SIGCHI Bulletin**, v. 23, n. 1, p. 82-87, jan. 1991.

NAUDTS, Laurens; DEWITTE, Pierre; ALVARADO RODRIGUEZ, Oscar; LAMBRECHT, Ingrid; MONTEIRO-KREBS, Luciana; STORMS, Elias; GEERTS, David; KUCZERAWY, Aleksandra; AUSLOOS, Jef. Algorithmic Transparency and Accountability in Practice (ATAP): A Study into Automated News Recommender Systems The Impact of News Recommender Systems. Conference: Interdisciplinariteit in Actie. KU Leuven. Leuven, Belgium. Jan. 2020. Disponível em: https://www.researchgate.net/publication/339747504_Algorithmic_Transparency_and_Accountability_in_Practice_ATAP_A_Study_into_Automated_News_Recommender_Systems_The_Impact_of_News_Recommender_Systems. Acesso em 18 out. 2024.

OBAR, J. A.; OELDORF-HIRSCH, A. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. **Information, Communication & Society**, v. 23, n. 1, p. 128-147, 2 jan. 2020.

PARKER, D. M.; PINE, S. G.; ERNST, Z. W. Privacy and informed consent for research in the age of big data. **Penn St. L. Rev.**, v. 123, p. 703, 2018.

PRAHALAD, C. K.; RAMASWAMY, V. **The future of competition**: Co-creating unique value with customers. Boston: Harvard Business Press, 2004.

SADOWSKI, J. When data is capital: Datafication, accumulation, and extraction. **Big Data & Society**, v. 6, n. 1, p. 1-12, 7 jan. 2019.

SUMÁRIO

- SANDERS, E. B. N.; STAPPERS, P. J. Co-creation and the new landscapes of design. **Co-design**, v. 4, n. 1, p. 5-18, 2008.
- SENDRA, P. The ethics of co-design. **Journal of Urban Design**, v. 29, n. 1, p. 4-22, 2024.
- SOUTHWICK, D.; RESCH, G.; RATTO, M. Iterative prototyping and co-design. In: **Knowledge, Innovation, and Impact: A Guide for the Engaged Health Researcher**. Cham: Springer International Publishing, 2020. p. 231-238.
- STORMS, ELIAS; ALVARADO, OSCAR; MONTEIRO-KREBS, Luciana . 'Transparency is Meant for Control' and Vice Versa: Learning from Co-designing and Evaluating Algorithmic News Recommenders. **Proceedings of the ACM on Human-Computer Interaction**, v. 6, p. 1-24, 2022. Disponível em: <https://dl.acm.org/doi/10.1145/3555130>. Acesso em: 18 dez. 2024.
- THATCHER, J.; O'SULLIVAN, D.; MAHMOUDI, D. Data colonialism through accumulation by dispossession: New metaphors for daily data. **Environment and Planning D: Society and Space**, v. 34, n. 6, p. 990-1006, 2016.
- TOUBIANA, V.; NISSENBAUM, H. **An analysis of google log retention policies**. 2011.
- UTZ, C. *et al.* (Un)informed consent: Studying GDPR consent notices in the field. In: **Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security**. New York, NY: ACM, 6 nov. 2019.
- VAN DEN BERG, B.; VAN DER HOF, S. What happens to my data? A novel approach to informing users of data processing practices. **First Monday**, v. 17, n. 7, 2012.
- VAN EVEN, P. **Trainer Deck Cards**. A reflection tool for the design, iteration, and evaluation of science stories. 2024. doi: 10.5281/zenodo.7589281.
- VAN EVEN, P. **Transcending the horizon of public science dissemination**. A foundational philosophical reflection on the science communication paradigm. Leuven: KU Leuven, 2023. ISBN 9789081428019.
- VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 3.4**. Evaluation report on science communication guidelines. Lappeenranta: Finland, 2022.
- VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 4.3 Trainer Training Package**. Lappeenranta, Finland, 2022.
- VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 3.6**. Final evaluation report. Lappeenranta, Finland, 2022.

SUMÁRIO

VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 4.4.** ParCos Trainer improvement report. Lappeenranta, Finland, 2022.

VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 4.2.** Report on science communication guidelines for the design and curation of participatory science stories. Lappeenranta, Finland, 2021.

VAN EVEN, P.; ZAMAN, B.; HANNES, K. **ParCos Deliverable 3.2.** Protocol for a systematic review on evaluation criteria for creative and interactive dissemination practices. 2020.

VAN OERS, R.; DE HART, B.; GROENENDIJK, K. **Country report:** the Netherlands. [s.l.: s.n.].

WALKER, S. C. *et al.* Strategies for enacting health policy codesign: a scoping review and direction for research. **Implementation Science**, v. 18, n. 1, p. 44, 2023.

WEBB, R. *et al.* Sustainable urban systems: Co-design and framing for transformation. **Ambio**, v. 47, p. 57-77, 2018.

WILLSON, Michele. Algorithms (and the) everyday. **Information Communication and Society** v. 20, n. 1, p. 137-150, 2017. Disponível em: <https://doi.org/10.1080/1369118X.2016.1200645>. Acesso em 18 jan. 2025.

WONG, C. C.; KUMPULAINEN, K.; KAJAMAA, A. Collaborative creativity among education professionals in a co-design workshop: A multidimensional analysis. **Thinking Skills and Creativity**, v. 42, p. 100971, 2021.

ZAMIR, E. Contract law and theory: Three views of the cathedral. **U. Chi. L. Rev.**, v. 81, p. 2077, 2014.

SUMÁRIO

SOBRE OS AUTORES E AS AUTORAS

Ana Julia Lopes

Doutoranda em Ciência da Informação na Universidade Federal do Rio Grande do Sul (UFRGS). Mestre em Ciência da Informação (UFRGS). Graduada em Bacharel em Biblioteconomia pela Universidade Federal do Rio Grande (FURG). Tem interesse nos eixos: Gestão de dados, Competência em dados, Repositórios digitais, Dados de pesquisa, Ciência Aberta.

Lattes: <http://lattes.cnpq.br/3811620567816952>

ORCID: 0000-0001-8552-5209

E-mail: lopesanajuliaci@gmail.com

Ana Maria Mielniczuk de Moura

Possui graduação em Biblioteconomia pela Universidade Federal do Rio Grande do Sul - UFRGS (1992), Especialização em Informática na Educação pelo PGIE/UFRGS em 2001, mestrado em Comunicação e Informação pelo PPGCOM/UFRGS em 2001 e doutorado em Comunicação e Informação pelo PPGCOM/UFRGS em 2009. Atualmente é professora Titular no Departamento de Ciências da Informação da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul e professora permanente no Programa de Pós-graduação em Ciência da Informação (PPGCIN/UFRGS). Coordena o Núcleo de Estudos em Ciência, Inovação e Tecnologia-NECIT desde 2017 e participa como membro do Grupo de Pesquisa em Comunicação Científica. Atua principalmente nos seguintes temas: ciência da informação, comunicação científica, informação científica e tecnológica, interação entre C,T &I, estudos métricos da informação e análise de patentes.

Lattes: <http://lattes.cnpq.br/1734997653639992>

ORCID: <https://orcid.org/0000-0002-7811-4660>

E-mail: ana.moura@ufrgs.br

SUMÁRIO

Carolina Carvalho Sena

Possui graduação em Biblioteconomia e Gestão de Unidades de Informação pela Universidade Federal do Rio de Janeiro (2010) e MBA em Gestão de Pessoas no Setor Público pela Universidade Estácio de Sá (2014). Mestre em Memória e Acervos pela Fundação Casa de Rui Barbosa (2018). Doutoranda em Comunicação, pelo Programa de Pós-Graduação em Comunicação, da Universidade Federal do Rio Grande do Sul (PPGCOM/UFRGS). É servidora pública federal, lotada atualmente na Seção de Disseminação de Informações, do Instituto Brasileiro de Geografia e Estatística (SDI/IBGE/RS).

Lattes: <http://lattes.cnpq.br/9396230860299406>

ORCID: <https://orcid.org/0009-0008-3146-6985>

E-mail: carolcarvalhosenaccs@gmail.com

Caterina Marta Groposo Pavão

Professora adjunta da Universidade Federal do Rio Grande do Sul. Tem experiência na área de Ciência da Informação, com ênfase em Biblioteconomia. Membro do Grupo de Pesquisa de Comunicação Científica da UFRGS e do Núcleo de Estudos em Ciência, Inovação e Tecnologia (NECIT). Coordenadora da Rede Sul de Repositórios Digitais. Possui graduação em Biblioteconomia e Documentação pela Universidade Federal do Rio Grande do Sul (1986), mestrado em Comunicação e Informação pela UFRGS (2010) e doutorado em Comunicação e Informação pela UFRGS (2014).

Lattes: <http://lattes.cnpq.br/4834791532698069>

ORCID: [0000-0003-3712-7200](https://orcid.org/0000-0003-3712-7200)

E-mail: caterina@ufrgs.br

Elisângela Gomes

Doutora (2023) e Mestra (2019) em Comunicação pela UFG; pós-graduada (lato sensu) em Teoria e Prática da Formação do Leitor pela UERGS (2015). Graduada em Biblioteconomia (UFRGS, 2012) e em Letras (ISEED, 2018). Pesquisadora do Núcleo Narrativas da Diferença (Pindoba) e do Laboratório Alaye (informação antirracista). Coordena o Clube de Leitura Maré Preta, é treineira do Grupo Calunga de Capoeira Angola e fundadora do Coletivo Pretas Angoleiras-GO. Atua em leitura e competência em informação, mediações culturais, metodologias decoloniais, raça, gênero e território, acervos visuais e bibliográficos, educação, literatura negra, narrativas literárias e memória.

Lattes: <http://lattes.cnpq.br/9006660307436079>

ORCID: <https://orcid.org/0009-0009-4308-0091>

E-mail: zanzamaio.go@gmail.com

SUMÁRIO

Fabiano Couto Corrêa da Silva

Docente da Universidade Federal do Rio Grande do Sul (UFRGS) e líder do Grupo de Pesquisa DataLab: Laboratório de Dados, Métricas Institucionais e Reprodutibilidade Científica. Atua nos cursos de graduação em Biblioteconomia (presencial e a distância), no Programa de Pós-Graduação em Ciência da Informação da UFRGS (PGCIN) e como colaborador do PGCIN/UFSC. Tem como principais temas de interesse: Gestão de Dados Científicos, Capitalismo de Vigilância, Colonialismo Digital, Indicadores e métricas de produção científica, Aprendizagem Colaborativa e Ciência Aberta.

Lattes: <http://lattes.cnpq.br/4635807083312321>

ORCID: <https://orcid.org/0000-0001-5014-8853>

E-mail: fabianocc@gmail.com

Fernanda Bochi

Doutora em Ciência da Informação do Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista (UNESP). Mestre em Comunicação e Informação pelo Programa de Pós-graduação em Comunicação e Informação e Bacharel em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (UFRGS).

Lattes: <http://lattes.cnpq.br/8572535542111719>

ORCID: <https://orcid.org/0000-0002-3275-0725>

E-mail: fernanda.bochi@unesp.br

Geise Ribeiro da Silva

Possui graduação em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (UFRGS), especialização em Práticas Pedagógicas pelo Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo (IFES), mestrado em Ciência da Informação pela UFRGS e é doutoranda em Comunicação na UFRGS. Interessa-se pelos temas: competência em informação, comportamento de busca e uso de informação, comunicação científica, divulgação científica e tecnologias de informação e comunicação.

Lattes: <http://lattes.cnpq.br/3447515096632830>

ORCID: <https://orcid.org/0000-0003-1658-6408>

E-mail: geisers2022@gmail.com

SUMÁRIO

Gláucia Aparecida Vaz

Professora Adjunta na Universidade Federal do Rio Grande do Sul/Faculdade de Biblioteconomia e Comunicação. Bolsista Júnior do CNPQ. Pós-doutorado em andamento na Universidade Estadual de Londrina. Doutora em Ciência da Informação, Mestre em Ciência da Informação e Bacharel em Biblioteconomia pela Universidade Federal de Minas Gerais, possui experiência em Biblioteca escolar, organização de acervos bibliográficos, fotográficos e documentais. Tem interesse em pesquisas sobre estudos de usuários e práticas informacionais em arquivos. Atualmente desenvolve estudos e palestras sobre leitura, literatura, representatividade negra na mídia e políticas educacionais antirracistas. Atualmente coordena projetos para implementação da lei 10.639/2003 em escolas públicas de educação básica. Como docente atua nas áreas de educação, leitura e literatura, leitura e sociedade, estudos de usuários em arquivos e bibliotecas e Metodologia da pesquisa científica.

Lattes: <http://lattes.cnpq.br/3514849083614884>

ORCID: <https://orcid.org/0000-0002-0249-5807>

E-mail: glauucia-vaz@hotmail.com

Jackson da Silva Medeiros

Professor do Programa de Pós-Graduação em Ciência da Informação e do Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul. Doutor em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul. Mestre em Ciência da Informação pela Universidade Federal Fluminense. Bacharel em Filosofia pelo Centro Universitário da Grande Dourados. Bacharel em Ciência Política pelo Centro Universitário Internacional. Bacharel em Biblioteconomia pela Universidade Federal do Rio Grande. Investiga como a informação estrutura a cultura, a sociedade e constitui o sujeito por meio de dinâmicas de poder, especialmente em contextos digitais.

Lattes: <http://lattes.cnpq.br/4182663628298542>

ORCID: <https://orcid.org/0000-0001-6725-4660>

E-mail: jacksonmedeiros@outlook.com

Janaina Lais Pacheco Lara Morandin

Doutoranda no Programa de Pós-graduação em Ciência da Informação da Universidade Estadual Paulista "Júlio de Mesquita Filho", mestre em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (UFRGS).

Lattes: <http://lattes.cnpq.br/8844607097133585>

ORCID: <https://orcid.org/0000-0001-7172-3087>

E-mail: janaina.morandin@unesp.br

SUMÁRIO

Jonas Ferrigolo Melo

Doutorando em Informação e Comunicação em Plataformas Digitais pela Universidade do Porto/Universidade de Aveiro (Portugal). Mestre em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul (2020). Especialista em História, Patrimônio Cultural e Identidades pela Universidade Luterana do Brasil (2014). Graduado em Arquivologia pela Universidade Federal de Santa Maria (2009). Arquivista do Arquivo Público do Estado do Rio Grande do Sul/APERS (2010 a 2022). Conselheiro do Conselho Nacional de Arquivos - CONARQ - Ministério da Justiça/Arquivo Nacional (2018 a 2022); membro da diretoria da Associação dos Arquivistas do Rio Grande do Sul (2015 a 2021). Membro do NUAWEB Núcleo de Pesquisa em Arquivamento da Web e Preservação Digital da UFRGS/ DGP e CNPq; investigador do CITCEM Centro de Investigação Transdisciplinar Cultura, Espaço e Memória / Faculdade de Letras da Universidade do Porto; e doutorando visitante do Meaningful Interactions Lab (Mintlab), University of Leuven (KU Leuven), Bélgica.

Lattes: <http://lattes.cnpq.br/0594057350445795>

ORCID: <https://orcid.org/0000-0002-7312-3509>

E-mail: jonasferrigolo@gmail.com

Jussara Borges

Possui graduação em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (1997), mestrado em Ciência da Informação (2005) e doutorado em Comunicação e Cultura Contemporâneas (2011), ambos pela Universidade Federal da Bahia. Realizou estágio doutoral na Universidade de Aveiro, Portugal (2010) e pós-doutoral na Universidade Carlos III de Madrid (2016) e na Universidad de Huelva (2024). É professora permanente do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN-UFRGS). Pesquisa e tem publicações da área de Informação e Comunicação, com ênfase em competências infocomunicacionais e desinformação. Foi bolsista de pós-doutorado no exterior pelo CNPq entre 2015 e 2016. Atualmente é coordenadora do PPGCIN e líder do Grupo de Pesquisa em Competências Infocomunicacionais (InfoCom).

Lattes: <http://lattes.cnpq.br/0229801641242896>

ORCID: <https://orcid.org/0000-0003-0157-8691>

E-mail: jussara.borges@ufrgs.br

SUMÁRIO

Letícia Pereira de Souza

Doutoranda em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC). Mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação na Universidade Federal do Rio Grande do Sul. (UFRGS). Graduada no curso de Biblioteconomia na Universidade Federal do Rio Grande do Sul (UFRGS). Técnica em biblioteconomia pelo Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS).

Lattes: <http://lattes.cnpq.br/1872733281003557>

ORCID: <https://orcid.org/0009-0002-6129-2337>

E-mail: leticiasouza.rs@hotmail.com

Lucia Helena Cunha Vidal

Doutoranda em Ciência da Informação (PPGCIN/UFRGS), Mestre em Ciência da Informação (PPGCIN/UFRGS) e graduada em Biblioteconomia (UFRGS). Pesquisadora em gestão de dados de pesquisa, planejamento e implementação de planos de gestão de dados.

Lattes: <http://lattes.cnpq.br/5256245097675300>

ORCID: [0009-0000-1431-5351](https://orcid.org/0009-0000-1431-5351)

E-mail: luciahcvidal@hotmail.com

Luciana Monteiro-Krebs

Professora no Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul (UFRGS, Brasil). Eleita uma das 100 Mulheres Brilhantes na Ética em Inteligência Artificial (100 Brilliant Women in AI Ethics 2025). Luciana também é membro do grupo de pesquisa DataLab - Laboratório de Dados, Métricas Institucionais e Reprodutibilidade Científica da UFRGS. Atuou como curadora de dados no projeto internacional da The Global Health Network (TGHN) na Fundação Oswaldo Cruz (Fiocruz), em colaboração com instituições de pesquisa na Europa, África, Ásia e América Latina (ex. Universidade de Oxford e HDR UK) (Hub TGHN da Fiocruz). Ainda na Fiocruz, Luciana integrou a equipe de apoio e curadoria de dados do Repositório oficial de dados para pesquisa da Fiocruz, o Arca dados. Anteriormente, no Meaningful Interactions Lab - Mintlab (KU Leuven), Luciana atuou em projetos de pesquisa em Interação Humano-Computador, como o 'Algorithmic Mediation in Academic Social Systems' (AMASS) e o 'Algorithmic Transparency Accountability in Practice' (ATAP), ambos projetos multidisciplinares e internacionais, pesquisando os efeitos dos algoritmos de recomendação na mídia a partir de uma abordagem centrada no usuário.

SUMÁRIO

Também trabalhou como cientista de dados no Cappra Institute (Brasil). Luciana é PhD em Ciências Sociais (KU Leuven, Bélgica) e Doutora em Ciência da Informação e Comunicação (UFRGS, Brasil). É especialista em organização do conhecimento, e é coautora do livro "Termos sob a superfície: Elementos teóricos, metodológicos e terminológicos para a representação do conhecimento" (Editora Interciência, 2021). Seus tópicos de interesse são ciência de dados, dados abertos de pesquisa, algoritmos de recomendação, comunicação científica, gamificação, padrões de uso em mídias sociais online, folksonomias, ontologias, análise de logs, recuperação da informação e terminologia.

Lattes: <http://lattes.cnpq.br/8151372917111773>

ORCID: <https://orcid.org/0000-0002-2882-6803>

E-mail: luciana.monteiro@ufrgs.br

Luis Fernando Massoni

Professor Adjunto do DFPI/ICI/UFBA e Professor Permanente do Programa de Pós-Graduação em Museologia e Patrimônio (PPGMusPa) da Universidade Federal do Rio Grande do Sul (UFRGS).

Lattes: <http://lattes.cnpq.br/1991208472616105>

ORCID: <https://orcid.org/0000-0001-6402-1036>

E-mail: luisfernomassoni@gmail.com

Maria Cláudia Cabrini Grácio

Livre-Docente em Estudos Métricos da Informação (2018) pela Universidade Estadual Paulista (UNESP), Doutora em Lógica (1999) e Mestre em Estatística (1991) pela Universidade Estadual de Campinas (UNICAMP). É docente do Departamento e da Pós-Graduação em Ciência da Informação da UNESP. Pesquisa no campo dos Estudos Métricos da Informação, principalmente nos campos: Bibliometria, Patentometria e Altimetria. É líder do Grupo de Pesquisa de Estudos Métricos da Informação.

Lattes: <http://lattes.cnpq.br/5170688300970006>

ORCID: <https://orcid.org/0000-0002-8003-0386>

E-mail: cabrini.gracio@unesp.br

SUMÁRIO

Moisés Rockembach

Professor do Programa de Pós-Graduação em Ciência da Informação (PPGCIN) e Programa de Pós-Graduação em Comunicação (PPGCOM), ambos da Universidade Federal do Rio Grande do Sul. Professor na Universidade de Coimbra. Doutor em Informação e Comunicação em Plataformas Digitais pela Universidade do Porto/ Universidade de Aveiro (Portugal), aprovado com distinção, título revalidado em 2014 no Brasil como Doutor em Comunicação e Informação (PPGCOM/UFRGS). Mestre em Comunicação e Informação pelo Programa de Pós Graduação em Comunicação e Informação (PPGCOM) da Universidade Federal do Rio Grande do Sul. Professor visitante na Universidade KU Keuven / Bélgica. Pesquisador do Núcleo de Pesquisa em Arquivamento da Web e Preservação Digital na Universidade Federal do Rio Grande do Sul, pesquisador do Centro de Estudos Interdisciplinares - Grupo Humanidades Digitais (CEIS20) e do projeto Inter paresTrust AI (University of British Columbia). Tem projetos de pesquisa financiados pelo CNPq (Edital universal 2021 - Chamada CNPq/MCTI/FNDCT No 18/2021), UFRGS, FAPERGS, RNP/IBICT e Social Sciences and Humanities Research Council of Canada (SSHRC). Membro da Camara Técnica Consultiva para a preservação de websites e mídias sociais do Conselho Nacional de Arquivos (CONARQ). Atuou como membro e parecerista da Comissão de Pesquisa e como Chefe do Departamento de Ciências da Informação da Faculdade de Biblioteconomia e Comunicação da UFRGS. Organização do XXII Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ENANCIB) na UFRGS. Membro de corpo editorial, editor convidado e parecerista de revistas científicas das áreas de Ciência da Informação e Comunicação no Brasil e no exterior. Autor de diversos artigos e livros, incluindo "Arquivamento da Web e Preservação Digital" (Pimenta Cultura, 2024), obra de referência no tema, em co-autoria com a Professora Doutora Caterina Groposo Pavão.

Lattes: <http://lattes.cnpq.br/1304688580274983>

ORCID: <https://orcid.org/0000-0001-9057-0602>

E-mail: moises.rockembach@gmail.com

Nicole Tirello Acquolini

Doutoranda em Comunicação pelo Programa de Pós-graduação em Comunicação (PPGCOM) da Universidade Federal do Rio Grande do Sul (UFRGS), mestrta em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação (PPGCIN/ UFRGS) e bacharela em Biblioteconomia pela mesma Universidade.

Lattes: <http://lattes.cnpq.br/8545707400153773>

ORCID: <https://orcid.org/0000-0002-3451-737X>

E-mail: nacquolini@gmail.com

SUMÁRIO

Pedro Henrique da Silva Rodrigues

Possui graduação em Biblioteconomia pela Universidade Federal do Rio Grande (FURG), mestrado em Ciência da Informação pela UFRGS e é doutorando em Ciência da Informação na UFRGS. Interessa-se pelos temas: comunicação científica, bibliometria, cientometria, estudos de citação, preprints e ciência aberta. Integra o Grupo de Pesquisa em Comunicação Científica (FABICO/UFRGS).

Lattes: <http://lattes.cnpq.br/7630185885673458>

ORCID: <https://orcid.org/0000-0001-5909-2067>

E-mail: pedrohenrique@ufrgs.br

Priscilla Van Even

Pesquisadora de pós-doutorado em ciência comunicação no Meaningful Interactions Lab. Estuda a tradução e a representação do conhecimento por meio de imagens e arte no projeto Palettes. Anteriormente, atuou como pesquisadora de doutorado no projeto ParCos, voltado ao apoio de práticas significativas de comunicação científica, e como pesquisadora em Interação Humano-Computador (IHC) no projeto RETINA, explorando como as tecnologias digitais podem apoiar as práticas de letramento visual de jovens em museus de arte e design. Além disso, lecionou metodologia de pesquisa na KU Leuven, possui experiência na educação de jovens e adultos, e atua como educadora de museu no Museu de Instrumentos Musicais e no Museu Real de Arte e História em Bruxelas. Priscilla é doutora em Sociologia, possui um Mestrado Avançado em Antropologia Cultural e Estudos do Desenvolvimento, um Mestrado em Filosofia, um Mestrado em Educação em Filosofia e Ciências Sociais, e um Bacharelado Profissional em Educação Infantil. Seus campos de especialização e interesses de pesquisa incluem comunicação científica, tradução do conhecimento, investigação paradigmática, mediatização da ciência, metodologia de pesquisa, biopolítica, patrimônio cultural e museus, saberes indígenas, saberes artesanais, filosofia da ciência, letramento visual, e filosofia e cultura orientais.

ORCID: <https://orcid.org/0000-0002-8272-1041>

E-mail: priscilla.vaneven@kuleuven.be

Rafael Port da Rocha

Possui graduação em Ciências de Computação pela Universidade Federal do Rio Grande do Sul (1988), mestrado em Computação pela Universidade Federal do Rio Grande do Sul (1992) e doutorado em Computação pela Universidade Federal do Rio Grande do Sul (2000). Atualmente é professor associado do Departamento de Ciência da Informação

SUMÁRIO

da Universidade Federal do Rio Grande do Sul. Tem experiência nas áreas de Ciência da Informação e Ciência da Computação, atuando principalmente nos seguintes temas: metadados, banco de dados, ontologias, preservação digital, gestão de dados da pesquisa. É fundador e atualmente coordena o Centro de Documentação e Acervo Digital da Pesquisa (CEDAP), órgão auxiliar da FABIO/UFRGS que busca dar suporte para pesquisas científicas, tecnológicas, artísticas e culturais realizada na Universidade, visando à gestão, preservação e uso científico e cultural de seus ativos digitais de pesquisa, natos ou decorrentes da digitalização

Lattes: <http://lattes.cnpq.br/5118387541734094>

ORCID: <https://orcid.org/0000-0003-4131-0534>

E-mail: rafael.rocha@ufrgs.br

Rene Faustino Gabriel Junior

Professor adjunto da Universidade Federal do Rio Grande do Sul. Tem experiência na área de Ciência da Informação, com ênfase em Biblioteconomia, atuando principalmente nos seguintes temas: bibliometria, brapci, ciência da informação, comunicação científica e produção científica. Implantou e coordena a Base de Dados de Periódicos em Ciência da Informação (BRAPCI). Membro do Grupo de Pesquisa de Comunicação Científica da UFRGS e do Núcleo de Estudos em Ciência, Inovação e Tecnologia (NECIT). Possui graduação em Biblioteconomia e Documentação pela Pontifícia Universidade Católica do Paraná (2008), mestrado em Ciência, Gestão e Tecnologia da Informação pela Universidade Federal do Paraná (2011) e doutorado em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (2014).

Lattes: <http://lattes.cnpq.br/5900345665779424>

ORCID: <https://orcid.org/0000-0003-1021-3360>

E-mail: rene.gabriel@ufrgs.br

Rodrigo Silva Caxias de Sousa

Professor Associado III do Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul (UFRGS). Bacharel em Biblioteconomia pela Universidade Federal do Rio Grande do Sul. (2000). Mestre em Educação pela Universidade de Passo Fundo (2003). Doutor em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul (2011). Atuou em bibliotecas escolares, universitárias e centros de informação no período anterior ao seu ingresso como professor do Departamento de Ciência da Informação. Também atuou como consultor em Informação e Documentação na empresa GLT Informação e Documentação, tendo experiência como consultor em

SUMÁRIO

editoras, bibliotecas comunitárias, bibliotecas escolares e bibliotecas universitárias. Como bibliotecário atuou nos setores de multimeios, referência e processamento técnico. Atualmente, como pesquisador, dedica-se ao estudos, que observam às seguintes ênfases: a) condicionantes do comportamento informacional na produção, comunicação e divulgação científica - processos de comunicação científica entre cientistas público leigo na web; percepção pública da ciência. b) desinformação e violências nas perspectivas de gênero e raça.

ORCID: <https://orcid.org/0000-0001-6872-4054>

E-mail: rodrigo.caxias@ufrgs.br

Samile Andréa de Souza Vanz

Professora titular do Departamento de Ciências da Informação. Professora permanente do Programa de Pós-graduação em Comunicação (PPGCOM UFRGS) e do Programa de Pós-graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN UFRGS). Graduada em Biblioteconomia pela Universidade Federal do Rio Grande do Sul (1999), mestre e doutora em Comunicação e Informação pelo PPGCOM UFRGS (2004 e 2009), com estágio sanduíche na Dalian University of Technology (China, 2007-2008). Pós-doutorado pela Universidad Carlos III de Madrid (Espanha, 2016). Desenvolve pesquisas na área de Comunicação Científica, com ênfase na produção de indicadores científicos, bibliometria, colaboração científica, análise de citação, análise de co-citação e rankings universitários. Tem experiência acadêmica e profissional na área de Planejamento, gestão e arquitetura de bibliotecas. Bolsista Produtividade PQ 2 (2016 -).ul.

Lattes: <http://lattes.cnpq.br/5243732207004083>

ORCID: <https://orcid.org/0000-0003-0549-4567>

E-mail: samilevanz@terra.com.br; samile.vanz@ufrgs.br

Sônia Elisa Caregnato

Professora titular da Universidade Federal do Rio Grande do Sul (UFRGS), integra o corpo docente do Programa de Pós-Graduação em Comunicação (PPGCOM), do Programa de Pós-Graduação em Ciência da Informação (PPGCIN) e do curso de graduação em Biblioteconomia. Lidera o Grupo de Pesquisa em Comunicação Científica e desenvolve estudos em comunicação científica, bibliometria, periódicos científicos, produção científica, repositórios digitais e acesso aberto a dados de pesquisa.

Lattes: <https://lattes.cnpq.br/5627209208288722>

ORCID: <https://orcid.org/0000-0002-5676-2763>

E-mail: sonia.caregnato@ufrgs.br

SUMÁRIO

Thaís Dias Medeiros

Doutoranda em Comunicação do Programa de Pós-Graduação em Comunicação (PPGCOM) da Universidade Federal do Rio Grande do Sul (UFRGS), Mestre em Comunicação no PPGCOM/UFRGS, Graduada em Biblioteconomia na Faculdade de Biblioteconomia e Comunicação da UFRGS.

Lattes: <http://lattes.cnpq.br/8540690399046551>

ORCID: <https://orcid.org/0000-0002-0104-1176>

E-mail: tmedeiros497@gmail.com

Thiago Henrique Bragato Barros

Professor Adjunto no Departamento de Ciência da Informação da Universidade Federal do Rio Grande do Sul (UFRGS), Professor permanente nos Programas de Pós-Graduação em Ciência da Informação da UFRGS e UFSC. Pós-Doutorado em Applied Informatics na Napier University (DES/CNPQ 2025-26) Pós-Doutorado em Ciência da Informação pela Universidade de Brasília (2018-2019). Pesquisador em grupos de pesquisa da UFSC, UFSCAR e UFRGS. Presidente da International Society for Knowledge Organization (2022-2026). Editor Chefe da revista em Questão (E-ISSN 1808-5245) Editor Managing Editor da revista Knowledge Organization (SSN 0943-7444) Coordenador do Programa de Pós-Graduação em Ciência da Informação - PPGCIN UFRGS(2020-2024). Vice- Presidente da Camara de Pós-Graduação CEPE/UFRGS (2024-2026). Vice Diretor da FABICO (2025-2028) Membro da Camara de Pesquisa do CEPE/UFRGS(2020-2024). Membro do Scientific and Tecnical Advisory Council da (STAC) da ISKO. Membro Associado da Association for Information Science and Technology. Membro do Conselho Internacional de Arquivos. Coordenou Grupo de Trabalho da Ancib. Presidiu a ISKO Brasil. Coordenou Curso de Graduação de Arquivologia e Dirigiu Faculdade de Arquivologia na UFPA. Organizou eventos nacionais e internacionais na área de informação. Foi representante da área da Informação no Comitê de Assessoramento na FAPERGS (2022-2025). Tem experiência na área de Ciência da Informação, atuando principalmente nos seguintes temas: Organização do Conhecimento, Análise de Domínio, Sistemas de Organização do Conhecimento e Arquivologia

Lattes: <http://lattes.cnpq.br/0339496971217162>

ORCID: <https://orcid.org/0000-0001-7439-5779>

E-mail: sean.vogel@gmail.com

SUMÁRIO

Valdir José Morigi

Professor titular aposentado do Departamento de Ciência da Informação/FABICO da Universidade Federal do Rio Grande do Sul (UFRGS). Possui graduação em Ciências Sociais pela Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS (1984), graduação em Biblioteconomia pela Universidade Federal da Paraíba - UFPB (1998), mestrado em Sociologia Rural pela Universidade Federal do Rio Grande do Sul (1988) e doutorado em Sociologia pela Universidade de São Paulo - USP (2001). Pós-doutorado em Memória Social pela UNIRIO (2017).

Lattes: <http://lattes.cnpq.br/6542370154854198>

ORCID: <https://orcid.org/0000-0002-2304-399X>

E-mail: valdir.morigi2@gmail.com

Vanessa Inácio de Souza

Possui graduação em Biblioteconomia (1999) e Mestrado em Ciência da Informação (2022) pela Universidade Federal do Rio Grande do Sul. É bibliotecária documentalista da mesma universidade. Tem experiência em processamento técnico de material bibliográfico, indexação, serviço de referência e comunicação na Web. Atua com capacitação de usuários em competências infocomunicacionais e é coordenadora adjunta do Curso de Extensão Super 8 - Pesquisa e Uso da Informação Científica.

Lattes: <http://lattes.cnpq.br/4393455153275129>

ORCID: <https://orcid.org/0000-0001-8542-7289>

E-mail: vanessa@bc.ufrgs.br

SUMÁRIO

ÍNDICE REMISSIVO

A

acesso aberto 21, 228, 234, 239, 313, 315, 317, 318, 321, 334, 335, 348, 349, 351, 389, 422, 482
acesso à informação 152, 279, 327
algoritmo 228
altmetria 22, 228
análise de conteúdo 26, 192, 198, 209
análise de domínio 91, 111
análise episódica 51, 52, 53, 55, 59, 61, 62, 63, 66
análise estatística 24, 393

B

bibliometria 22, 26, 179, 181, 205, 251, 254, 256, 274, 275, 277, 280, 479, 481, 482
big data 21, 145, 399, 414, 415, 422, 423, 469

C

capitalismo 57, 451
capitalismo de vigilância 451
ciência aberta 12, 14, 15, 17, 21, 23, 260, 312, 315, 321, 326, 328, 415, 479
ciência cidadã 315, 318, 319, 336, 340, 341, 347, 464
ciência da informação 10, 11, 12, 13, 37, 69, 88, 89, 143, 162, 218, 281, 312, 313, 353, 432, 472, 481
cientometria 22, 479
coautoria 21, 250, 256, 257, 258, 259, 267, 268, 272, 274, 281, 330, 343, 350, 392
co-design 13, 17, 311, 432, 433, 434, 435, 436, 437, 438, 439, 444, 445, 450, 451, 452, 453, 454, 455, 457, 458, 460, 464, 465, 466, 467, 468, 470, 471
colonialidade 34, 41, 54, 57, 60
colonialismo de dados 451
competências infocomunicacionais 16, 117, 119, 129, 130, 134, 135, 136, 137, 142, 476, 484

computação 15, 18, 89, 162, 421

comunicação científica 11, 17, 18, 21, 91, 100, 185, 220, 221, 223, 224, 227, 234, 244, 254, 260, 281, 322, 386, 438, 439, 440, 443, 444, 472, 474, 478, 479, 480, 481, 482

consentimento informado 453

curadoria digital 23, 337, 357, 375, 376, 428

D

dados intensivos 21
dados pessoais 446, 447, 448, 450, 453, 467
dados sensíveis 316
decolonialidade 32, 34, 53, 60
democratização do conhecimento 319, 320, 321, 332
descolonização 52, 57, 58
desobediência epistêmica 33, 34, 42, 61, 64

E

epistemicídio 58, 60
epistemologia aplicada 16
epistemologias decoloniais 32, 34
escrevivência 35, 36, 37, 38, 39, 40, 42, 65
espaços periféricos 34
estudos de usuários 26, 465, 474, 475
estudos métricos da informação 162, 179, 251, 472
experiência 32, 36, 38, 42, 52, 53, 56, 62, 64, 67, 79, 85, 121, 127, 269, 285, 286, 452, 455, 473, 474, 480, 481, 482, 483, 484

G

gênero 37, 41, 43, 55, 56, 123, 295, 301, 304, 307, 481
governança de dados 383, 400, 401, 415, 421

I

imagens de controle 43, 48, 50, 51, 66

SUMÁRIO

indexação 16, 91, 97, 103, 105, 107, 170, 184, 222, 227, 228, 239, 255, 259, 484

informetria 22

infraestruturas digitais 106

integridade científica 415

inteligência artificial 145, 183, 228, 234, 269, 279, 280

J

justiça social 199

L

largura 385

LGPD 446, 447, 453, 467

linguagens documentárias 22, 91

livros 43, 45, 168, 203, 228, 229, 234, 237, 238, 256, 264, 266, 267, 294, 352, 369, 479

M

mediação da informação 22, 137

memória social 12, 16, 284, 285, 286, 294, 298, 306, 307

metodologias de pesquisa 14, 199, 285, 325, 384

métodos mistos 24, 25

N

negritude 57, 58

neurociência 325

O

ontologias 21, 90, 91, 93, 97, 98, 99, 104, 108, 109, 376, 478, 480

open access 351

open data 427

open peer review 21, 352

P

paradigmas de pesquisa 18, 19

patentometria 16

pesquisa qualitativa 196, 329

positivismo 41, 45, 118

práticas pedagógicas antirracistas 37

preservação digital 12, 17, 320, 336, 337, 344, 350, 354, 355, 356, 357, 358, 361, 363, 364, 365, 366, 367, 368, 369, 370, 371, 373, 374, 375, 376, 378, 480

princípios FAIR 323, 324, 344, 390, 396, 397, 415, 417, 420, 423

processo científico 331

produção científica 20, 22, 23, 24, 55, 115, 167, 183, 187, 188, 199, 201, 202, 206, 207, 212, 215, 226, 251, 252, 253, 258, 259, 260, 262, 263, 265, 267, 278, 279, 282, 314, 316, 335, 342, 348, 351, 474, 481, 482

R

raça 41, 43, 44, 47, 56, 59, 63, 64, 66, 481

racismo cotidiano 28, 51, 52, 56, 63, 68

repositórios institucionais 24, 237, 321

representação do conhecimento 21, 27, 477, 480

revisão sistemática 19, 261, 262, 277, 407, 414, 429, 439

S

sexismo 42, 43, 48, 66

T

teoria psicanalítica 52

transparência metodológica 21

U

uso da informação 22, 78, 130, 151, 254

www.PIMENTACULTURAL.com

CAMINHOS METODOLÓGICOS EM CIÊNCIA DA INFORMAÇÃO

fundamentos
e aplicações



CAPES



PPGCIN
Programa de Pós-Graduação
em Ciência da Informação



pimenta
cultural