

DISCRIMINAÇÃO POR DADOS: UMA ANÁLISE A PARTIR DA LITERATURA DE CIÊNCIA DA INFORMAÇÃO

José Augusto Bagatini, Universidade Estadual Paulista (Unesp), <https://orcid.org/0000-0002-8830-2075>

Suellen Oliveira Milani, Universidade Federal Fluminense (UFF), <https://orcid.org/0000-0001-7183-6030>

**José Augusto Chaves Guimarães, Universidade Estadual Paulista (Unesp),
<https://orcid.org/0000-0002-0310-2331>**

RESUMO

As representações usadas para organizar o mundo vão além da organização do conhecimento produzido, e, cotidianamente, envolve classificar pessoas. Entretanto, nota-se que a classificação de pessoas pode resultar em situações excludentes e, em última instância, à perseguição. Isso posto, a problemática de pesquisa decorre do ato de nomear pessoas segundo uma lógica capitalista em que o que importa é como esse indivíduo pode consumir e colaborar para o acúmulo de capital. Dessa forma, e considerando a atualidade e a relevância científica e social da discriminação por dados, o objetivo dessa pesquisa é identificar como a literatura da área de Ciência da Informação aborda a referida temática, mais especificamente suas abordagens, ênfases e desdobramentos. Para tanto, foram realizadas buscas nas bases de dados Library and Information Science Abstracts (LISA) e Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), contemplando tanto o contexto internacional quanto o brasileiro. O corpus obtido foi submetido ao método de análise de domínio, o qual visa a identificar e categorizar um universo de conteúdo a partir de seus contextos de produção. Identificou-se que o tema é novo e demanda pesquisas colaborativas e interdisciplinares. Conclui-se que as discussões se situam predominantemente em veículos que abordam práticas profissionais da Ciência da Informação, incluindo discussões arquivísticas ou acerca da ética informacional. A preocupação da literatura em caracterizar traços distintivos desse tema de forma agregadora revela que ele pode atuar como um intertema investigativo em distintas áreas.

Palavras-Chave: Organização do Conhecimento; Proteção de Dados; Dados Pessoais; Discriminação por Dados; Vieses na Representação.

DISCRIMINACIÓN POR DATOS: UN ANÁLISIS DESDE LA LITERATURA DE LAS CIENCIAS DE LA INFORMACIÓN

RESUMEN

Las representaciones utilizadas para organizar el mundo van más allá de la organización del conocimiento producido y, cotidianamente, implican clasificar a las personas. Este problema de investigación surge del acto de nombrar a las personas según una lógica capitalista en la que lo importante es cómo este individuo puede consumir y colaborar para la acumulación de capital. De esta forma y considerando la actualidad y la relevancia científica y social de la discriminación de datos, el objetivo es identificar cómo la literatura de las Ciencias de la Información entiende este tema, especialmente en cuanto a sus enfoques, énfasis y ramificaciones. Para eso, se realizaron investigaciones en las bases de datos Library and Information Science Abstracts (LISA) y Reference Database for Journal Articles in Information Science (BRAPCI), abarcando tanto el contexto internacional como el brasileño. El corpus obtenido fue sometido al método de análisis de dominio,

que tiene como objetivo identificar y categorizar un universo de contenidos a partir de sus contextos de producción. El tema es nuevo y exige un trabajo colaborativo e interdisciplinario. Se concluye que las discusiones se publican predominantemente en revistas que abordan las prácticas profesionales, incluidas las discusiones de archivo y las discusiones sobre ética de la información. La preocupación de la literatura en caracterizar rasgos distintivos de este tema de forma agregada revela que puede actuar como intertema investigativo en diferentes áreas.

Palabras-Clave: Organización del Conocimiento; Protección de Datos; Datos Personales; Discriminación por Datos; Sesgos en la Representación.

DATA DISCRIMINATION: AN ANALYSIS FROM THE INFORMATION SCIENCE LITERATURE

ABSTRACT

Representations used to organize the world go beyond the organization of the knowledge produced, and, daily, involve classifying people. This research problem stems from the act of naming people according to a capitalist logic in which what matters is how this individual can consume and collaborate for the accumulation of capital. In this way and considering the actuality and the scientific and social relevance of data discrimination, the goal is to identify how Information Science literature understand this theme, especially about its approaches, emphases, and ramifications. For that, searches were carried out in the Library and Information Science Abstracts (LISA) and Reference Database for Journal Articles in Information Science (BRAPCI) databases, covering both the international and Brazilian contexts. The corpus obtained was submitted to the domain analysis method, which aims to identify and categorize a universe of content from its production contexts. The topic is new and demands collaborative and interdisciplinary work. It is concluded that the discussions are predominantly published in journals that address professional practices, including archival discussions and discussions about information ethics. The concern of the literature in characterizing distinctive features of this theme in an aggregating way reveals that it can act as an investigative inter-topic in different areas.

Keywords: Knowledge Organization; Data Protection; Personal Data; Data Discrimination; Bias in Representation.

1 INTRODUÇÃO

A organização do conhecimento acompanha a história, como uma necessidade social para fazer frente ao conhecimento que se produz e que é socializado para gerar novo conhecimento, em um movimento helicoidal contínuo (Guimarães, 2008). Nesse contexto, Pombo (1998) refere-se à tríade da organização: dos seres, dos saberes e dos documentos, respectivamente com a classificação de Lineu, no século XVII e todas as taxonomias desde então existentes; a trajetória que vai de Aristóteles, passando por Bacon, Harris, e os enciclopedistas; e a tradição da Biblioteconomia e da Documentação com Dewey, Brown, Otlet, Ranganathan e outros (San Segundo, 1996). Em

toda essa trajetória percebe-se a preocupação com a construção de sistemas de organização do conhecimento.

Como toda atividade humana, a organização do conhecimento, por vezes lida com os vieses sociais dos responsáveis pelas representações (catalogadores, indexadores etc.). Olson (2002), enfatiza o fato de essas representações atuarem, para a comunidade, como substitutos dos documentos do conhecimento (surrogates of knowledge), gerando instrumentos mediadores entre as necessidades, desejos e demandas de informação dos usuários e os recursos informacionais presentes nos acervos

disponibilizados. Quando essa representação não é realizada de uma forma defensável, podem ocorrer distorções (censuras, omissões etc.) com o risco de se privar alguém de ser ouvido ou de receber informações (Olson, 2002; Milani & Guimarães, 2017).

Essa representação, por sua vez, faz-se por meio de um controle imposto à linguagem, baseado não nas características e necessidades de cada indivíduo, mas em um padrão que ampare uma comunidade de usuários. Nesse contexto, corre-se o risco de a representação gerar tendenciosidade evocando as crenças e valores éticos do profissional, ou ainda, veiculando preconceitos, tendenciosidades ou vieses – biases – sejam decorrentes da ação de quem realiza a classificação ou a indexação, ou de quem construiu um sistema de organização do conhecimento. Com isso, corre-se o risco de ocorrerem efeitos danosos aos usuários, principalmente aos que não pertencem às comunidades dominantes (Olson, 2002; Milani & Guimarães, 2017).

Invocando Dale Spender, Olson (2002) chama a atenção para o fato de que toda

2 CLASSIFICAR DADOS SOBRE PESSOAS

A organização do mundo ao nosso redor vai além do conhecimento produzido, e, cotidianamente, envolve classificar pessoas, de maneira quase inconsciente e para sanar objetivos sociais. Veja-se, para tanto, o censo demográfico, que remonta a dois milênios antes de Cristo, na China, como atividade básica para o Estado colher informações importantes sobre sua população (IBGE, 2012).

Desde então, as tecnologias empregadas nos censos evoluíram e, no censo estadunidense de 1890 utilizou-se pela primeira vez o aparato das máquinas elétricas capazes de analisar dados coletados em papéis perfurados, idealizadas por Hollerith cuja empresa viria a tornar-se, em 1924, a *International Business Machines Corporation* (IBM). Essa tecnologia, por sua vez, foi utilizada pelo governo nazista de Adolf Hitler e, para Black (2001), o holocausto

nomeação é necessariamente tendenciosa pois efetiva um processo de codificação de vieses (no simples fato de se eleger uma representação em detrimento de outra). Essa nomeação está diretamente conectada à linguagem que, possui elementos de natureza subjetiva e política de tal modo que o poder de nomear molda e define as instituições que estruturam grande parte de nossas vidas, indo além da mera representação para construir uma informação.

Isso posto, a problemática de pesquisa decorre do ato de nomear pessoas segundo uma lógica capitalista em que o que importa é como esse indivíduo poderá consumir e colaborar para o acúmulo de capital. Dessa forma, e considerando a atualidade e a relevância científica e social da discriminação por dados, tem-se por objetivo identificar como a literatura da Ciência da Informação vem tratando a referida temática, em especial no que se refere a suas abordagens, ênfases e desdobramentos.

só pôde ser tão eficiente porque a tecnologia da IBM forneceu subsídios suficientes para a realização de um censo rápido e eficaz, em que se pôde identificar e catalogar os judeus na Alemanha, cujos dados sobre ele e seus familiares eram registrados em cartões perfurados de cores diferentes, o que garantiu a construção de uma base de dados ampla e rica em informações sobre a população judia alemã. Com a informação em mãos, o regime nazista conseguiu gerenciar o processo de expropriação de direitos e bens de judeus. As máquinas da IBM também foram utilizadas nos campos de concentração para o processo de gerenciamento de presos e de suas atividades, inclusive o fornecimento de alimentos aos prisioneiros em níveis mínimos para a sobrevivência e a execução de trabalhos

forçados, mantendo-os em seu limite físico, de modo que não pudessem oferecer resistência.

Nota-se que a classificação de pessoas, quando levada ao extremo, pode resultar em atividades excludentes e, em última instância, levar à perseguição. Entretanto, após a Segunda Guerra Mundial e, principalmente após a popularização da internet, a coleta de dados sobre pessoas foi amplificada, e não mais passou a ser atividade restrita a governos.

Bauman e Lyon (2012), chamam a atenção para o aparecimento e consolidação daquilo que denominam economia da privacidade, a qual é composta por instituições que possuem seus modelos de negócios voltados à vigilância e que orientam seus negócios e produtos para a coleta e processamentos de dados pessoais, estabelecendo, assim, uma faceta do capitalismo que se fundamenta no enfraquecimento do direito à privacidade. Somado a isso, Mayer-Schönberger (2011) afirma que a erosão da privacidade individual é um dos principais desafios a ser enfrentado na contemporaneidade uma vez que os dados coletados se tornam cada vez mais abrangentes, registrados e mantidos indefinidamente em uma memória digital que pertence a companhias e governos (Mayer-Schönberger, 2011).

Essa memória, a qual Mayer-Schönberger (2011) refere-se, não se contenta apenas com dados demográficos, como na primeira metade do século XX. Governos e empresas atuam hoje como “buracos negros de dados”, absorvem, armazenam e processam todo e qualquer tipo de dado de proveniência pessoal.

Atualmente é possível ler notícias, consumir filmes e séries, conectar-se com amigos, fazer compras, acompanhar dados de saúde como sono e ciclo menstrual, tudo através de dispositivos como smartphones, relógios e pulseiras inteligentes. Essa trilha de dados sobre uma pessoa, foi batizada por Alan Westin como *data shadow*, ainda na década de

1960, em sua obra “Privacy and Freedom”. Segundo Saulles (2015), esses conjuntos de dados, podem ser divididos em estruturados e não estruturados, o primeiro originado a partir de transações em que uma das partes é a instituição que coleta os dados, como na compra e venda de uma mercadoria, enquanto o segundo grupo diz respeito a dados produzidos a partir de outros tipos de interação, como e-mails, telefonemas, publicações em redes sociais etc. (Saulles, 2015).

Esses dados de natureza pessoal vêm sendo explorados de diversas maneiras por empresas e governos, entre suas possíveis utilizações, servem como insumos para algoritmos que geram análises e decisões automatizadas, as quais podem conter vieses que aprofundam assimetrias no tratamento que um grupo social pode receber.

O Conselho de Estado Chinês publicou, em 2014, um documento no qual prevê a criação e implementação de um plano de crédito social. Através desse programa, o governo chinês pretende mapear, através de dados oriundos de diversas fontes, condutas sociais, financeiras e comportamento tanto online quanto offline. Cada ação poderá ser categorizada como “aprovada” ou “reprovada”, resultando assim, em soma ou subtração de pontos atribuídos aos cidadãos do país (Tumelero, 2019). Caso o cidadão perca demasiados pontos, o acesso a serviços de transporte como trens e aviões, serviços de internet, acesso a hotéis de melhor qualidade, escolas e universidades melhores, vagas de emprego podem ser revogados e, pode até mesmo chegar em situações extremas na qual o nome do cidadão pode ser inserido em listas públicas e mensagens de voz serem exibidas quando uma pessoa liga para o telefone do cidadão mal posicionado no ranking (Tumelero, 2019).

Em Wisconsin, é utilizado um algoritmo chamado COMPAS, sigla para Correctional Offender Management Profiling for Alternative Sanctions, cujo objetivo é avaliar a probabilidade de um indivíduo voltar a cometer

um crime, e através disso, determinar uma pontuação que define a pena. Os dados utilizados pelo sistema são amplos e compreendem desde histórico de prisão de familiares até desempenho escolar do analisado, além disso, o software aplica uma série de questões ao acusado e julga se há ou não “pensamentos criminosos”. O funcionamento do programa é secreto, segredo comercial, ou seja, é impossível contestar uma pontuação, afinal, não se pode saber como ela foi obtida. Em uma investigação, a ProPublica, corporação sem fins lucrativos de jornalismo investigativo, identificou que minorias étnicas recebem pontuações maiores a partir de uma análise comparativa entre os acusados de cor preta e branca com a mesma idade, sexo e ficha criminal. O resultado demonstrou que acusados pretos possuem 45% mais chances do que brancos de receber uma pontuação alta (Instituto Humanitas Unisinos, 2016).

3 PROCEDIMENTOS METODOLÓGICOS

Essa pesquisa exploratória partir de um levantamento bibliográfico nas bases de dados *Library and Information Science Abstracts* (LISA) e Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), contemplando tanto publicações internacionais quanto brasileiras. No caso da LISA, as estratégias de busca foram “*algorithm* bias* and 'personal data**”, b) “*data discrimination**”, com os refinamentos de texto completo e artigo de revista. Foi recuperado um total de 13 artigos na literatura internacional. Na base BRAPCI, a partir dos termos “*vie** de *algorithm**” and “*dados*

4 RESULTADOS

Observou-se, no *corpus* da pesquisa, um conjunto de 27 autores diferentes, cada qual figurando apenas uma única vez, o que revela que o tema é ainda novo – o primeiro artigo foi publicado em 2016, e dois terços

Em 2015, um usuário do Google denunciou que o algoritmo da plataforma havia rotulado a foto de um casal negro com uma legenda contendo o nome de um animal (g*****). Esse mesmo erro aconteceu com o Flickr, serviço de armazenamento e publicação de fotografias quando a plataforma categorizou as fotos de pessoas negras com a tag “m*****” (Cruz, 2021).

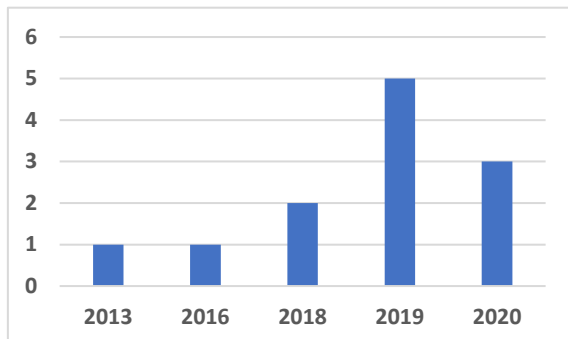
A vista do exposto, fica evidente que na atual sociedade digital, os algoritmos presentes em diversas atividades geram injustiças porque baseiam-se em modelos matemáticos que podem reproduzir preconceitos, equívocos e vieses humanos. Esses modelos são enviesados porque existe alguém que decide o que é importante ou não, e geralmente essas ferramentas são construídas em função da maximização dos lucros, não havendo incentivos para tornar os sistemas mais justos, o objetivo ético não costuma acompanhar o dinheiro (O’Neil, 2016).

“*personais**” e discriminação por dados”, sem nenhum resultado.

Dos treze documentos recuperados foram eliminados quatro por não serem artigos, por não oferecerem metadados descritivos que possibilitassem uma análise de seu conteúdo, ou por não se referirem ao tema. Chegou-se, assim, a um *corpus* de nove artigos ao qual aplicou-se o método de análise de domínio, (Hjørland & Albrechtsen, 1995; Hjørland, 2002, 2017), que visa a identificar e categorizar um universo de conteúdo a partir de seus contextos de produção.

foram publicados entre 2019 e 2020 – sem tempo para a sedimentação de referentes teóricos.

Gráfico 1: Artigos publicados por ano.



Fonte: Elaborado pelos autores (2022).

Essa produção revelou uma natureza eminentemente colaborativa, com média de 3 autores por artigo, o que demonstra que embora o tema seja recente, já se delineiam equipes de pesquisa.

Tabela 1: Lista de autores

Autores	Artigos
Alexios Mantzarlis	1
Assunta Hunter	1
Cathy Pitkin	1
Christine O'Keefe	1
Debora Warr	1
Derina Holtzhausen	1
Elizabeth Englezos	1
Filippo Menczer	1
Gabriele Spina Ali	1
Giovanni Luca Ciampaglia	1
Gregory Maus	1
Jenny Waycott	1
Jodie Mcvernon	1
Joelle B. Thorpe	1
Jolene Jesse	1
Kerin Clark	1
Luyue Ma	1
Malkiat Thiarai	1
Marilyn Guillemín	1
Matt Duckham	1
Priscilla M. Regan	1
Richard Sinnott	1
Ronald Yu	1
Sarunkorn Chotvijit	1
Sherry Wasilow	1

Stephen Jarvis	1
Steven Prawer	1
Total	27

Fonte: Elaborado pelos autores (2022).

Esse conjunto de autores provém de quinze diferentes vinculações institucionais, com destaque para as universidades de Melbourne (7 autores), de Indiana e Warwick (3 autores cada), bem como a *Commonwealth Scientific and Industrial Research Organisation* e o Departamento de Defesa Nacional do Canadá (2 autores cada).

Tabela 2: Vinculações institucionais

Instituições	Autores
The Universit of Melborne	4
Indiana University	3
University of Warwick	3
Commonwealth Scientific and Industrial Research Organisation	2
Defence Research and Development Canada	2
Melbourne School of Population and Global Health	2
Babes-Bolyai University	1
Berks Thun Library	1
George Mason University	1
Griffith University	1
iBIONICS	1
Lamar University	1
Melbourne Law School	1
National Science Foundation	1
Poynter Institute	1
RMIT University	1
University of Washington	1
Total	27

Fonte: Elaborado pelos autores (2022).

Essa produção procede de tão somente cinco países: Austrália (11 autores), Estados Unidos (9 autores), Canadá (3 autores), Reino Unido (3 autores) e Romênia (1 autor). Ressalta-se que esse universo de produção científica provê, em sua franca maioria (95%) de países de língua inglesa.

A produção foi veiculada em sete periódicos distintos, sendo que em apenas dois

deles houve dois artigos: AI Magazine e Ethics and Information Technology, respectivamente voltados para a inteligência artificial e para o diálogo entre filosofia moral e as tecnologias em informação e comunicação.

Tabela 3: Títulos dos periódicos científicos

Periódicos	Artigos
AI Magazine	2
Ethics and Information Technology	2
Cambridge University Press	1
Erecords Management Journal	1
Journal of Communication	1
Journal of Information Ethics	1
Journal of Information, Communication and Ethics in Society	1
Total	9

Fonte: Elaborado pelos autores (2022).

As 30 palavras-chaves dos artigos (média de 3,3 por artigo), referem-se a:

- Atributos inerentes aos dados (12 ocorrências)
- Contextos de aplicação ou relações interdisciplinares do tema (8 ocorrências)
- Aspectos éticos (5 ocorrências)
- Objeto dos processos, contextos e problemas (4 ocorrências)
- Problemas (1 ocorrência)

A análise temática do *corpus* aponta para a preocupação com potenciais ameaças para a comunicação estratégica, pois os contextos sociais e culturais do *big data* se constroem a partir daqueles que criam algoritmos, sobrepondo seus valores e

5 CONSIDERAÇÕES FINAIS

Os preconceitos veiculados por representações causam impacto na autoimagem dos usuários, pois além de não cumprirem sua função como pontes de acesso ao conhecimento, reforçam exclusões sociais. Com isso, as comunidades de usuários que se

referências culturais aos dados (Holtzhausen, 2016).

A questão ética se revela relativamente a uso de inteligência artificial e de robótica pelas forças armadas canadenses para fins de defesa e segurança (Wasilow & Thorpe, 2019), à privacidade, anonimidade, monitoramento, não-discriminação e propriedade da informação na utilização de tecnologia educacionais (Regan & Jesse, 2019), à governança da informação (Thiarai et al., 2019) e às suas decorrências a partir da expansão do uso de dados digitais para formação de diretrizes para as necessidades de pesquisadores e comitês de ética, possibilitando avaliar os riscos de pesquisas que façam uso de dados pessoais (Clark et al., 2019).

Os problemas incidentes se referem a: *misinformation* e aos desafios de pesquisa para a construção de uma web confiável (Ciampaglia et al., 2018); proliferação dos dados pessoais pelo fato de a proteção à privacidade continuar se concentrando no consentimento no espaço digital (Englezos, 2020); e possibilidade de os sistemas de Inteligência Artificial escaparem do controle e compreensão de seus operadores e programadores, resultando em distorções ou falhas e comprometendo a confiabilidade da pesquisa jurídica (Yu & Ali, 2019).

A busca de possíveis soluções, de acordo com Ma (2020), envolve o fornecimento de uma infraestrutura interdisciplinar para melhor compreender o potencial democratizante das tecnologias digitais e identificar novas perspectivas acerca da opacidade e da confiabilidade dos algoritmos e de como isso pode afetar o potencial de democratização das tecnologias digitais.

encontram às margens, por não pertencerem às comunidades dominantes na sociedade, não são representadas de forma defensável e podem sofrer discriminação por misoginia, homofobia, xenofobia, intolerância racial e religiosa, preconceito de classe, descrição racial

equivocada, entre outros. Essas situações, seja em substitutos documentais seja na decisão automatizada e com base em dados pessoais, atingem questões de acesso à informação, essencial para diminuição de desigualdades decorrentes de um modelo econômico capitalista.

A análise do *corpus* revelou que o tema é ainda novo e vem demandando pesquisas

colaborativas e interdisciplinares. As discussões da temática se situam predominantemente em veículos que abordam práticas profissionais da Ciência da Informação, incluindo discussões arquivísticas ou de ética informacional. Revela ainda a preocupação da literatura em caracterizar traços distintivos nessa temática de forma agregadora, que pode atuar como um intertema investigativo em distintas áreas.

REFERÊNCIAS

- Bauman, Z. & Lyon, D. (2012). *Liquid Surveillance: A Conversation*. Polity. <https://papers.ssrn.com/abstract=3791056>.
- Black, E. (2001). *IBM and the holocaust: The strategic alliance between nazi germany and America's most powerful corporation*. Elsevier.
- Ciampaglia, G. L., Mantzarlis, A., Maus, G. & Menczer, F. (2018). Research Challenges of Digital Misinformation: Toward a Trustworthy Web. *AI Magazine*, 39(1), Art. 1. <https://doi.org/10.1609/aimag.v39i1.2783>.
- Clark, K., Duckham, M., Guillemin, M., Hunter, A., McVernon, J., O'Keefe, C., Pitkin, C., Praver, S., Sinnott, R., Warr, D. & Waycott, J. (2019). Advancing the ethical use of digital data in human research: Challenges and strategies to promote ethical practice. *Ethics and Information Technology*, 21(1), 59–73. <https://doi.org/10.1007/s10676-018-9490-4>.
- Cruz, B. S. (2021). *Efeito de algoritmos racistas pode ser mais devastador do que você imagina*. <https://www.uol.com.br/tilt/reportagens-especiais/como-os-algoritmos-espalham-racismo/>.
- Englezos, E. (2020). *Forget Consent? Answering the Challenges of Digital Space* (SSRN Scholarly Paper ID 3791056). Social Science Research Network.
- Guimarães, J. A. C. (2008). A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). *Revista Ibero-Americana de Ciência da Informação*, 1, 77–79.
- Hjørland, B. (2002). Domain analysis in information science. Eleven approaches—Traditional as well as innovative. *Journal of Documentation*, 58(4), 422–462.
- Hjørland, B. (2017). Domain analysis. *Knowledge Organization*, 44(6), 436–464.
- Hjørland, B. & Albrechtsen, H. (1995). Toward a new horizon in information science: Domain analysis. *Journal of the American Society for Information Science*, 46(6), 400–425.
- Holtzhausen, D. (2016). Datafication: Threat or opportunity for communication in the public sphere? *Journal of Communication Management*, 20(1), 21–36. <https://doi.org/10.1108/JCOM-12-2014-0082>.
- IBGE (2012). *O censo tem história*. <https://cnae.ibge.gov.br/en/component/content/article/290-teen/censo/censo-2000/1703-o-censo-tem-historia.html>.

- Instituto Humanitas Unisinos. (2016). *Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA*. <http://www.ihu.unisinos.br/78-noticias/561869-sistema-de-algoritmo-que-determina-pena-de-condenados-cria-polemica-nos-eua>.
- Ma, L. (2020). Rethinking democratizing potential of digital technology: A review of technology and communication studies. *Journal of Information, Communication and Ethics in Society*, 18(1), 140–156. <https://doi.org/10.1108/JICES-02-2019-0022>.
- Mayer-Schönberger, V. (2011). Delete: The virtue of forgetting in the digital age. Em *The demise of forgetting—And its drivers* (4^o ed). Princeton University Press.
- Milani, S. O. & Guimarães, J. A. C. (2017). Problemas relacionados a biases em Sistemas de Organização do Conhecimento: Perspectivas para a <https://doi.org/10.1007/s10676-018-9492-2>.
- San Segundo, R. (1996). *Sistemas de Organización del Conocimiento*. Madrid: *Bolentin Oficial del Estado; Universidad Carlos III de Madrid*.
- Saulles, M. de. (2015). *Information 2.0: New models of information production, distribution and consumption* (Vol. 2). Facet.
- Thiarai, M. Chotvijit, S., & Jarvis, S. (2019). Balancing information governance obligations when accessing social care data for collaborative research. *Records Management Journal*, 29. <https://doi.org/10.1108/RMJ-09-2018-0029>.
- Tumelero, T. (2019). *Crédito Social: A realidade experimentada da China*. NSC total. <https://www.nsctotal.com.br/noticias/credito-social-a-realidade-experimentada-da-china>.
- representação de assunto. Em F. A. Pinho & J. A. C. Guimarães (Orgs.), *Memória, tecnologia e cultura na Organização do Conhecimento* (p. 399–407). Ed. UFPE.
- Olson, H. A. (2002). *The power to name: Locating the limits of subject representation in libraries*. Kluwer.
- O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Pombo, O. (1998). Da classificação dos seres à classificação dos saberes. *Leituras, Revista da Biblioteca Nacional de Lisboa*, 2, 19–33.
- Regan, P. M. & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: Twenty-first century student sorting and tracking. *Ethics and Information Technology*, 21(3), 167–179.
- Wasilow, S. & Thorpe, J. B. (2019). Artificial Intelligence, Robotics, Ethics, and the Military: A Canadian Perspective. *AI Magazine*, 40(1), Art. 1. <https://doi.org/10.1609/aimag.v40i1.2848>.
- Yu, R. & Ali, G. S. (2019). What’s Inside the Black Box? AI Challenges for Lawyers and Researchers. *Legal Information Management*, 19(1), 2–13. <https://doi.org/10.1017/S1472669619000021>.