Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization: A Thematic Domain-Analysis

José Augusto Bagatini and* José Augusto Chaves Guimarães**

* ** São Paulo State University – UNESP, Av. Hygino Muzzy Filho 737 – 17525-900 – Marília – SP – Brazil
*<jose.bagatini@unesp.br> / **<chaves.guimaraes@unesp.br>

José Augusto Bagatini holds a Bachelor degree in Archival Sciences and a Master degree in Information Science. He is currently a PhD student in Information Science at the Graduate School of Information Science of São Paulo State University and holds a research fellowship from the Brazilian Coordination of Superior Level Staff Improvement – Capes. His research interests are: Privacy, Personal Data Protection and Digital Surveillance, especially within the Information Science field.

José Augusto Chaves Guimarães is Professor Emeritus at the Graduate School of Information Science of São Paulo State University – UNESP, and also a collaborator professor at the Graduate School of Information Science of Fluminense Federal University, in Brazil. He is the founder of the Brazilian chapter of ISKO, and a former member of the Executive Board of ISKO. His research interests are: knowledge organization (especially on cultural and ethical issues), domain analysis, scientific communication and legal information.

Abstract: Personal data play a fundamental role in contemporary socioeconomic dynamics, with one of its primary aspects being the potential to facilitate discriminatory situations. This situation impacts the knowledge organization field especially because it considers personal data as elements (facets) to categorize persons under an economic and sometimes discriminatory perspective. The research corpus was collected at Scopus and Web of Science until the end of 2021, under the terms "data discrimination", "algorithmic bias", "algorithmic discrimination" and "fair algorithms". The obtained results allowed to infer that the analyzed knowledge domain predominantly incorporates personal data, whether in its behavioral dimension or in the scope of the so-called sensitive data. These data are susceptible to the action of algorithms of different orders, such as relevance, filtering, predictive, social ranking, content recommendation and random classification. Such algorithms can have discriminatory biases in their programming related to gender, sexual orientation, race, nationality, religion, age, social class, socioeconomic profile, physical appearance, and political positioning.

1.0 Introduction

The protection of personal data is a relevant concern in the contemporary world and, especially in the field of Information Science, it presents challenges and perspectives in relation to the processes, instruments and products of knowledge organization and representation.

Personal data are generally understood to be those relating to a natural person, based on elements associated with their location or their physical, physiological, genetic, mental, economic, cultural, or social identity. Canadian legislation (Canada, 2000) emphasizes that such data may be factual or subjective, recorded or not, while the World Economic Forum includes not only data but also metadata created by and about a person, either voluntarily as in social network profiles, as a result of observation, when captured by different means, or even based on inferences when based on third-party analysis, as is the case with a financial score (World Economic Forum 2021).

Personal data, and more specifically sensitive data – such as racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union, membership, sex life, or sex-

ual orientation – can serve as subsidy for processes that aim at the segregation of a particular social group, such as the situation described by Edwin Black (2001) in which personal data, collected from a population census, served as input for the holocaust, enabling Nazi Germany to identify, catalog and persecute Jews, gypsies, blacks, and gays in the country. This leads to the phenomenon of data discrimination, whose segregating configuration brings consequences of an ethical nature.

Given the above, the objective of this paper is to conceptually delimit and thematically categorize data discrimination, based on the international literature, as this theme is especially relevant to knowledge organization and representation theories and practices.

2.0 The use of personal data in contemporary society and the ethical impacts on knowledge organization

Knowledge organization accompanies the history of humanity as a social need to face all the produced knowledge, which is socialized to generate new knowledge, in a continuous helical movement [1] (Guimarães 2008). In this sense, Pombo (1998) refers to a universe that incorporates the taxonomic dimension, which dates back to Linnaeus in the 17th century in relation to living beings and with repercussions until today; the classificatory dimension, which appears with Aristotle, passing through Bacon, Harris, and the encyclopedists in order to categorize knowledge; and the tradition of Librarianship and Documentation, aimed at rescuing document content, with Dewey, Brown, Otlet, Ranganathan and many others (San Segundo 1996). The concern with the construction of representation instruments is observed throughout this trajectory.

In a world permeated by technology, the practices of organizing and representing knowledge go beyond the documentary universe and often affect the identity of human beings because, just as in personal and private relationships, in which people are classified, organized and represented according with aspects such as kinship or level of intimacy, in the same logic, and on a larger scale and considering administrative objectives, governments have historically used census techniques to collect data from their population that allow a comprehensive view of their socioeconomic characteristics and, not rarely, they enter into important ethical questions. To illustrate this phenomenon, in the 1930s, and using IBM technology, the German government was able to carry out its population census with great speed and efficiency. During the data collection process, individuals who identified themselves as Jews had their personal and familial information recorded on perforated cards of a distinct color. This led to the creation of a comprehensive and extensive database on the Jewish population in Germany, enabling

the implementation of a plan to seize assets from the Jewish community. This technological use to capture and organize data even reached extreme situations when data was managed, including even on the prisoners′ diet, making calculations to keep the prisons at a minimum energy level for survival and forced labor, that is, they kept them in their physical limit so that they could not offer any resistance (Black 2001).

After World War II, several technologies originally created in a military context were adapted to collect data in different areas, generating concerns related to the so-called data shadow, that is, personal data recorded and maintained by companies and governments in everyday situations (Westin 1967), and which can be structured (such as those derived from of a purchase situation in a store) or unstructured (such as those resulting from interactions such as phone calls, for example (Saulles 2015).

The situation has become increasingly complex with the advent and widespread use of personal computing and internet communication, which has led to a significant shift in how we create and store our memory artifacts (captures, sharing, etc.) in order to ensure portability in space and permanence in time (Smit and Barreto 2002; Mayer-Schönberger 2011). As a result, the storage and representation of memory artifacts have evolved from being initially confined to blogs, then expanding to social networks, and finally reaching cloud storage services. This evolution has enabled memory artifacts to connect with private and public identities, giving rise to novel experiences and representations of oneself and others. This phenomenon triggered a process of self-erosion of anonymity, inducing us to make public what is private, and assuming the end of invisibility and anonymity as real, putting the right to privacy at risk (Van Dijck 2007; Bauman and Lyon 2012; Ketelaar 2014).

The issue of privacy has been escalating at geometric progression with the widespread adoption of technologies such as smartphones, wearables, and the Internet of Things. These technologies are generating an unprecedented volume of personal data, capturing diverse aspects of human life. Consequently, individuals are now carrying with them their own vulnerable personal surveillance devices, connected and transmitting data to an economy that prioritizes the commodification of personal data while compromising the right to privacy (Bauman and Lyon 2012).

For Westin (1967, 26), privacy has to do with the informative self-determination of individuals, groups or institutions characterized by the "voluntary and temporary withdrawal of a person from general society through physical or psychological means, either in a state of solitude or small-group intimacy or, when among larger groups, in a condition of anonymity or reserve." For the author, privacy performs four basic functions for individuals: personal autonomy (development and maintenance of a sense of indi-

338

Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

viduality, based on the desire to avoid being manipulated or dominated by others); emotional release (exercise of authenticity, outside the performance of social roles); self-evaluation (provision of time and space for the individual to carry out his "moral inventory"); and limited and protected communication (possibility of sharing confidences and intimacies with others who are trusted). Véliz (2021) even states that preserving privacy is a manifestation of power.

However, contemporary societies have challenged the privacy of individuals, notably from instruments of physical surveillance (optical or acoustic devices of the location, acts, speech or private writing of a person without their knowledge or against their will, psychological, oral or writing tests, and devices or substances to extract information from an individual when they are unaware of its disclosing), and data (collection, exchange and manipulation of document information about individuals and groups by data processing machines). The widespread adoption and rapid advancement of Information and Communication Technologies (ICTs) in generating knowledge and facilitating decision-making has led to an increased demand for data collection and record-keeping across all sectors of society. Thus, this increase in information collection and processing, not carefully regulated, led to a significant growth in the surveillance power of both governments and organizations over individual lives (Westin 1967).

It is worth noting that massive surveillance policies permeate the entire development of the Internet as we know it today and have contributed to defining guidelines on how the *World Wide Web* can provide subsidies for espionage actions, with access to data such as search history, complete emails, files, chats, etc., stored on servers that belong to *Google, Facebook, Apple* and other technology companies (Greenwald and MacAskill 2013). A quick analysis of *Google's* growth history since its foundation in 1998 reveals that its business model, aiming to expropriate personal data, is profitable. Its success stems from its search engine with a simple interface but with a robust operating system, currently indexing almost the entire web and considering each user's personal preferences. Thus, by combining data from more than 1.2 trillion searches per year (Saulles 2015) with other data such as login, cookies, IP addresses, etc., *Google* can accurately identify any person who has already used any of its services. In addition, the company holds a memory that details an entire human life, things the user itself had forgotten or discarded long before (Mayer-Schönberger 2011).

Offering free services that have as a bargaining chip the appropriation of users' personal data is not an exclusive practice of *Google*, constituting a standard of action in the privacy market. This is combined with the so-called data brokers, which operate almost exclusively in the capture, processing and sale of personal data, or even the insight ser-

vices that use this data, such as: *Acxiom, Corelogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf, and Recorded Future* (Federal Trade Commission 2014). These companies use what Acquisti et al.(2016) advocate as trade-offs in the privacy market and feed their personal databases from various sources, which capture data in situations such as: the purchase of a common good (in which the consumer can leave data during the financial transaction); data exchange between the agents of the privacy market; and the personal information protection market (when consumers look for products and services to manage and protect their personal data).

Trying to equalize the power relationship between the owner of personal data and those interested in exploiting these data, laws have emerged in several countries to regulate the privacy market, a phenomenon that has been closely monitored by the United Nations Conference on Trade and Development – UNCTAD, through its Data Protection and Privacy Legislation Worldwide page (UNCTAD 2020).

This scenario of constant expansion of the Internet, hyperconnectivity (Castells 2009) and the integration of humanity in the digital economy raises issues related to the most diverse fields of Information Science, such as the intersections between memory and the right to be forgotten, the protection of the citizen, and knowledge organization and the segregation of social minorities.

The actors that make up the privacy market have been establishing, as a marketing standard, memory bases in which the concept of temporality is distorted according to their interests, since the maintenance of personal data ensures the continuous production of inferences and, consequently, the prediction of future actions and behavior. These memory bases are generally inaccessible to the data subjects, which will be difficult to erase in the future, and they can be used against the population at any time. This leaves the question of how ethical it is for an institution to store in its digital memory something that compromises a person's present and future, directly affecting the right to be forgotten (Mayer-Schönberger 2011; Saulles 2015).

Thus, it is necessary to change the approach regarding the safeguarding of personal data through the adoption of a model in which: a) the life cycle stages of data based on the rights to privacy and to be forgotten; b) the collection/capture of personal data in a digital or physical environment, and in a manual or automated way, is based on ethical principles such as a legitimate reason, data subject (or legal guardian) consent; c) the collection methodology and data typology are duly specified and publicized to the data subject when subscribing to the service or product. To this end, data storage must occur in a safe environment (with well-defined preservation goals and criteria, anonymized and encrypted), the safeguard must occur for a limited period (until the objectives manifested at the beginning of the collec-

tion are achieved 1; the retrieval must be carried out through authorization and consider authentication tools that can generate trails and detailed logs that subsidize the rendering of accounts to the data subjects, and the disposal must occur due to the fulfillment of the objective or upon data subject's request. However, data may be transferred to another institution, due to institutional needs or at the user's request, in which case strong rules must be used to ensure a safe transfer and data integrity, and with equivalent institutional policies of the recipient similar in degree of protection and transfer, and duly informed to the data subject (Cavoukian 2006; Romansky 2015; Sant'Ana 2016; Poikola et al. 2020; Bagatini et al. 2021).

In order to mitigate the damages resulting from Internet users' activities, Masur (2020) proposes literacy in privacy, which goes hand in hand with digital literacy, based on the development of the critical capacity of individuals to decide for themselves, when and within what limits information about them should be collected, analyzed, stored and disseminated. This assumes the understanding of responsibilities and risks associated with sharing information online. Privacy literacy takes place through citizens[2] understanding of the information landscape with which they interact, the responsibilities that result from it, awareness of how information is tracked and used in online environments, and the possibilities to maintain or to lose the information's private nature. This presupposes knowledge about the technical aspects[3] of legal online data protection so that the individual can be properly informed about the storage, use and distribution of their personal data and decide which of these data are public or private in nature. Furthermore, this literacy should encompass five dimensions: understanding how personal information is used online, recognizing the places where personal information may be shared online, realizing the consequences of sharing personal information online, assessing the risks and benefits of sharing information, and deciding when to share this information online (Langenderfer and Miyazaki 2009; Bartsch and Dienlin 2016; Wissinger 2017).

This complex scenario affects knowledge organization, especially due to its potential for segregation, since knowledge organization focuses on establishing conceptual structures and representing them. These two activities are not neutral as they result from the contexts in which they are inserted and which are permeated with values. Olson (2002) referred to the 'power to name,' originally coined for libraries in the establishment of 'surrogates of knowledge' in catalogs, a situation that has expanded in scope and specificity in the digital reality, with the tags that guide algorithms in retrieval and use of personal data and which, similarly to catalogs – but with greater power – can promote discrimination and marginalization of certain individuals and groups. These surrogates of knowledge, when not con-

structed in a defensible way, can censor, omit and distort information, deprive someone of being heard or of receiving information. Thus, the 'power to name', when carried out in a biased manner, generates biases by evoking unilateral ethical beliefs and values and, as a consequence, generates harmful effects mainly to those who do not belong to mainstream communities (Milani and Guimarães 2017). Naming is directly connected to language, which is one of the most intimate and political activities, so the power of name shapes and defines the institutions that structure much of our lives. Naming information is not simply information representation, but also the construction of that information (Olson 2002).

In the context of knowledge organization, the scientific community has been studying, especially over the last two decades, the ethical challenges imposed on this field. In this sense, Beghtol (2002; 2005) warns of the need for a cultural warrant of representations so that the represented being identifies with its representation. This, as highlighted by Gutiérrez (2002), presupposes the adoption of a transcultural ethics of mediation that allows a dialogue between the represented being, its *representamen* and those who use the representation. In this context, (Guimarães 2017) highlights the importance of a representation that is attentive to local identities without disregarding the need for global dialogues. This ethics, in turn, must be based on values that offer a barrier to the dissemination of prejudices and antipathies (Berman 1993) as these are, by definition, segregating and excluding. Fernández-Molina et al. (2005) highlight, in turn, how these risks have become potentiated with the growth of ICTs.

As previously stated, humans are no exception as objects of naming, classification and representation processes and, currently, are constantly subjected to analysis by automated systems that, when using personal data, can generate decisions with biases and, thus, deepen asymmetries in the treatment that a group may receive. In other words, it can be said that individuals become, in this context, sources of knowledge that are analyzed from a set of data (classification facets) combined by algorithms to compose a representation (notation through tags) that will be used for its general categorization of services for marketing purposes and not rarely, discriminatory ones.

For O'Neil (2016), data discrimination deepens inequalities and threatens democracy since the computational algorithms present in various activities of society generate injustices because they are based on mathematical models that can reproduce prejudices, mistakes and human biases once they come from an instance that decides what is important or not, and generally these tools are built to maximize profits. In this regard, mention is made of the dismissal of an American teacher based on the algorithms of a scoring system to identify less productive teachers without knowing

340

Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

the source of the data (Menárguez 2018). There is also the case of the algorithm COMPAS – Correctional Offender Management Profiling for Alternative Sanctions (program is secret, in commercial secrecy) to assess the probability of an individual committing a crime again and, from there, determine a score that defines the penalty. In an investigation, ProPublica, a non-profit investigative journalism corporation, identified that ethnic minorities receive higher scores from a comparative analysis between black and white defendants with the same age, sex and criminal record (Instituto Humanitas Unisinos 2016).

Data discrimination is an aspect of capital importance today, with harmful and often irreparable consequences for the citizen who, in a marketing logic, is seen as a unit of knowledge whose data become the object of classification and representation, without their consent, for purposes that are not always defensible, often generating discrimination and segregation. This highlights the importance of data discrimination as a knowledge domain that interacts with knowledge organization and, therefore, needs to be better understood in its conceptual dimension, its characteristics and scope, based on the international scientific literature that has been addressing the topic.

3.0 Methodology

To achieve the proposed objectives, domain analysis was carried out (Hjørland and Albrechtsen 1995; Guimarães 2014; Hjørland 2017), based on the international literature on the subject of algorithmic discrimination/data discrimination, using two of the eleven approaches proposed by Hjørland (2002) – epistemological studies (under the thematic dimension) and bibliometric studies. To this end, and based on the two axes provided by Tennis (2003) to operationalize a domain analysis, the Scopus and Web of Science databases were used as areas of modulation (extension) by the end of 2021 and, as degrees of specialization, the search terms "data discrimination", "algorithmic bias", "algorithmic discrimination" and "fair algorithms" with the limiters "file type: article", "access: open", "publication stage: final" and "language: Spanish, English and Portuguese." Thus, in a preliminary corpus of 276 scientific articles, a refinement was carried out to eliminate record duplication in both databases, documents in formats other than articles, and documents that were not retrieved or were no longer available, resulting in a set of 187 articles.

The analysis of the corpus was carried out using the Content Analysis methodology, (Bardin, 2013), coming from the area of Sociology and widely used in Human and Social Sciences literature, in works of a qualitative nature and in themes that can, by their nature, be expressed with greater terminological variety, as is the case with the ethical aspects of a field of knowledge. In this way, content analysis is used

to reach, in greater depth, the thematic dimension of the corpus.

Content Analysis, as advocated by Laurence Bardin (2013), is carried out based on inferences drawn from the content of documents according to a categorization arising from them, which, according to Sales (2008), provides, at the same time, freedom to the analyst and objectivity of investigation. This method consists of pre-analysis phases (construction of an exhaustive, representative and pertinent analysis corpus and formulation of hypotheses based on this corpus based on a categorization process through the establishment of analysis categories with convergent characteristic features), and exploration of the material (from reading the corpus and extracting and describing information), and processing the results (from applying the categories to the material extracted from the corpus) (Guimarães and Sales 2010).

Thus, in the first phase there is the definition of the research corpus and what will be sought in it, using principles of exhaustiveness (the corpus in its entirety), representativeness (quality of the corpus, which is guaranteed by the reputation of the sources and due to the rigor in search strategies); homogeneity (in this case, all were articles from periodicals contained in the databases researched); and relevance: adequacy of the recovered corpus to the objectives of the analysis. In the second phase, based on the corpus, the analysis categories are established, while in the third phase, the results and interpretation of this analysis are presented.

Therefore, and considering the nature of the research corpus, the diversity of languages and the desired qualitative approach, it was decided to carry out the analysis process manually, although we are aware of the significant contribution of the NLP (Natural Language Processing) techniques by means of some widely available tools (such as NVivo 12) which will group items in a corpus of literature together based on a statistical analysis of the word frequency etc.

After reading the documents, it was found that, although presenting keywords relate to the researched topic, only 58 articles effectively addressed it and, therefore, represent the referred knowledge domain: data discrimination and the performance of algorithms incident to it.

At first, we sought to characterize this literature in terms of authors' origin, average life and publication vehicles, and then proceeded with content analysis (Bardin 2013) in order to reach, in greater depth, the thematic dimension of the corpus [4].

4.0 Results: presentation, analysis, and discussion

In order to identify the existence of a group of researchers addressing the subject, an analysis of the authors of the 58 articles that make up the final corpus was carried out, reach-

Knowl. Org. 50(2023)No.5

341

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

ing a total of 163 authors. In this context, a great dispersion of authors was observed since only one author – O. Papakyriakopoulos[5] – has published two articles, while all the others have published only one. The data shows that the theme has been the object of concern from various scientific communities, but it is still in its early stages of scientific production. The lack of repeated publications shows that no author has an investigative trajectory on this topic or a more significant core of authors representing the theme.

As for the author's institutional affiliation, 96 universities located on different continents were found, and it is noteworthy that only five universities published more than one article each: the European universities of Bern, Leiden, and London, and the US universities in Arizona and California. All other universities were responsible for only one article. It is also observed that the group of universities is basically from the northern hemisphere, mainly European and North American regions where the theme has been more traditionally treated, including in legal instruments.

Regarding the authors' analysis, it was verified that the theme is also pulverized institutionally, not forming a core of authors and institutions that constantly address the theme since only five of the ninety-six universities have two articles linked to them.

Regarding the countries of the universities where the authors are associated [6], it is worth mentioning that the United States, United Kingdom, Spain and Switzerland together account for 57% of the academic community that researches the topic, with more than a dozen researchers each. However, there is a reasonable number of researchers seeking to develop the theme in countries of the "socioeconomic south" [7], with emphasis on China, Indonesia, India and Brazil.

Regarding discrimination for purposes of national identity, the production of this literature basically took place from 2017 (with 3 articles), i.e., one year after the approval of the EU General Data Protection Regulation (GDPR). However, with the GDPR in effect in 2018 (3 articles), a strong growth of scientific production on the topic can be observed from the following year onwards, with 7 articles in 2019, 20 in 2020 and 24 in 2021.

It is essential to highlight the importance of GDPR in the face of increased scientific production, since it is a standard of significant global reach, which evokes explicit consent, promotes the consolidation and respect of individual rights concerning personal data (since it provides citizens with the ability to access, rectify, delete, port and restrict the processing of data relating to them); and constitutes an incentive for organizations to incorporate the notion of privacy by design from the beginning of the development of their processes and products. Such aspects, among other issues, directly impact different fields of knowledge (such as

Information Science), culminating in the growth of scientific interest in privacy and protection of personal data.

Regarding the publication vehicles of the analyzed corpus, a set of 54 journals was found, of which only three – AI & Society, Information, Communication and Society, and PlosOne (with respectively three, two and two articles each) – published more than one article on the subject. The thematic grouping of journals, based on the areas of knowledge to which they belong, as stated in their editorial scope, shows a strong concern of a multi, inter and transdisciplinary nature, with almost half of the universe, which reveals the transversal aspect of this theme, reaching different areas of knowledge in society. The second incidence occurs in journals in the field of Computer Science (20%), which was to be expected since issues related to artificial intelligence and the configuration of algorithms are traditionally studied in this area and, thirdly are the Communication and Information journals (11%), revealing a concern for clarifying the general public, alerting users, and providing information about risks and damages, as well as measures to avoid them. In a more isolated way, and showing a more applied treatment of the theme, there are journals of Law, Medicine, Administration and Psychology, among others. In the specific scope of Information Science, the articles were published in Information Processing & Management; International Information and Library Review; Proceedings of the Association for Information Science and Technology.

To characterize the thematic content of the research corpus, the keywords of the articles were first analyzed and then categorized and grouped through content analysis (Bardin 2013). Thus, it was possible to establish thematic categories of keywords. On a scale, there are keywords linked to the ethical aspect of the theme (94 incidences), to the very object of the analyzed domain – the data and the action of algorithms in the scope of artificial intelligence (74 incidences), and the discriminations and the biases they affect (45 incidences), the contexts and ambiances in which data discrimination occurs (62 incidences), the specific verified processes and procedures (42 incidences), and the theories, methodologies and interdisciplinarities that guided the approaches of the articles (31 incidents). Finally, two more specific universes were evident, albeit in a much less significant way: the explicit mention of behavior (5 incidences) and information (4 incidences) as permeating elements of the theme. The thematic categories of keywords and respective incidences are summarized below:

- **_Ethics (94 incidences):_** Ethics (AI, dating, technology); Ethical values: (Care, Equality, Explainability; Fairness, Identity, Interpretability, Justice, Non-discrimination, Privacy, Resilience, Trust); Ethical problems (Authoritarianism, Colonialism, Digital activism, Discrimination; Inequality, Popularity, Prejudice, Racism, Surveil-

342   Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

lance, Unfairness); Biases (algorithmic, classification, cognitive, confirmation, gender, racial, sample, survey).

– ***Data, Algorithms, and Artificial Intelligence (74 incidences):*** Algorithm (randomized, recomender); Algorithmic (stability, accountability, bias, colonisation, decision-making, discrimination, distribution, fairness, filtering, video editing,); Artificial intelligence (connectionist, emotional, ethics, responsible, symbolic); Data (big, imputation, personal, protection, regulation, standards, visualization).

– ***Contexts and ambiances (62 incidences):*** Academic libraries; Castes; Civil Society (Protests); Health (care, disparities, technology, behavioral, digital, mental, vaccination); Hyperactive users; LGBTQAI+; Networks (complex, neural, social); Places (Latin America, European Union, Russia, North Korea, Mexico, Lower-middle-income-countries); Pole dance; Policies (Technological policy); Politics (Elections, Political Communication, Political data Science, Political information, Political machines); Science (Open, Participatory); Search engines (Yandex, Google); Systems (Recommender systems, Smart information; System); Uber.

– ***Processes and Actions (42 incidences):*** Deep learning; Facial recognition and analysis; Human decision-making; Machine learning; Measurement; Modeling; Monitoring; Natural language processing; Networking; Profiling; Ranking; Regulation; Risk assessment; Shadowban; Shared decision-making; Signaling.

– ***Theories, Methodologies and Interdisciplinarities (31 incidences):*** Theories (Neo-platonism, Pragmatism, Realism, Standpoint theory); Methodologies (Agent-based testing, Bayesian analysis, Biometrics; Case studies, Counterculture Empirical studies, Cross-cultural perception, Demographic parity, Feature selection, Network analysis, Omitted variables; Predictive analytics, Randomized classification, Survey); Interdisciplinarities (Cybernetics, Design, Medicine, Computational Social Science, Culture, Philosophy of Science, Sociology).

– ***Behavior (05 incidences):*** Advertising, Health, Human, Human-computer interaction; Online.

– ***Information (04 incidences):*** Information (evaluation, literacy); Disinformation; Misinformation.

Following, and in order to retrieve possible approaches not represented in the universe of keywords, a content analysis of the titles and abstracts of the articles was carried out, since these are the textual parts in which the authors seek to describe the theme, the object and purpose of their articles in a synthetic way. Thus, and similarly to the content analysis carried out with the keywords, eight thematic categories were identified: a) types of discriminatory biases; b) ambiances of application and realization contexts; c) involved values and ethical problems; d) types of addressed algorithms;

e) types of involved data; f) tools and instruments designed and used; g) explicit mention of specific theories and methodologies; and h) interactions and consequences of the theme. The classification of the content along with the respective categories, subcategories and the articles that refer to them are described in Table 1.

It is important to highlight that the terms – "data discrimination", "algorithmic bias", "algorithmic discrimination", and "fair algorithms" were used for the purpose of retrieving literature from the thematic synthesis areas of the articles – title, abstract and /or keywords – and highlight elements of the area's terminology. On the other hand, the thematic categories and subcategories identified highlight conceptual spaces resulting from content analysis – which may even be – or not – terminologically represented in the documents through the four search terms used.

Other, more specific reflections underlying data discrimination were present, relating to power relations between companies, platforms and users (Rettberg 2020), hidden elements in content filtering technologies (Peralta et al. 2021), the importance of studying terminology and the nature of biases to neutralize algorithmic discriminations (Rozado 2020), the characterization of Data Science as a technological Neoplatonism insofar as algorithmic discrimination can transform predictions into a right of preference and the people involved in it run the risk of an abstraction of responsibility and generating "carelessness" (McQuillan 2018), the discussion of privacy as a contribution of the human sciences to artificial intelligence (Ostherr 2020), the need to carry out audits in algorithmic systems (Shen et al. 2021), and the importance of developing media literacy in users, and especially in librarians as a way to deal with algorithmic biases (Gardner 2019).

Finally, it is highlighted that, in an analysis carried out over a decade ago by Guimarães et al. (2008) regarding the scientific production published in the Journal of the American Society for Information Science and Technology — JASIST, Journal of Documentation, Knowledge Organization, Cataloging & Classification Quarterly, The Indexer, and Ethics and Information Technology, between 1995 and 2004 [8], 23 values [9] and 21 problems[10] of ethical nature related to the field of knowledge organization were identified. The present study revealed that seven of those values (Cultural warrant, Equity, Freedom, Information safety, Privacy, Reliability, and Risk minimization) and eleven of those problems (Censorship, Defamation, Digital segregation, Idiosyncrasy, Lack of cultural warrant, Marginalization, Misrepresentation, Prejudices in terminology, Racism, Surveillance, and Violence) are present in the context of algorithmic discrimination.

Category	Subcategory	Authors
Ethical values and problems	*Surveillance capitalism*	(Jooste 2021)
	Algorithmic Colonization	(Jooste 2021)
	Justice and algorithmic trust	(Rosenbaum and Fichman 2019) (Du et al. 2020) (Kristiadi et al. 2020) (Kehrenberg et al. 2020) (Paulus and Kent 2020) (Zhang et al. 2021) (Ostherr 2020) (Sun et al. 2020)
	Oppressive digital practices	(Jooste 2021)
	Algorithmic responsibility	(Rosenbaum and Fichman 2019)
	Algorithmic transparency	(Abul-Fottouh et al. 2020) (Bolander 2019) (Zhang et al. 2019)
	Interoperability in algorithmic decisions	(Bolander 2019) (Zhang et al. 2021)
Discriminatory biases	*Cognitive*	(Yan 2021)
	Physical appearance	(Rozado 2020)
	Shadowban	(Are 2021)
	Social class and socio-economic profile	(Schuler and Montardo 2020) (Verma and Acharya 2018) (Williams et al. 2018) (Evangelo and Oliveira 2021) (Rozado 2020) (Zajko 2021)
	Gender and sexual orientation	(Fosch-Villaronga et al. 2021) (Martínez-Plumed et al. 2021) (Schuler and Montardo 2020) (Khalil et al. 2020) (Gupta et al. 2021) (Rozado 2020)
	Age	(Akter et al. 2021) (Khalil et al. 2020)
	Nationality	(Martínez-Plumed et al. 2021) (Akter et al. 2021) (Castro 2020) (Yu et al. 2012)
	Political positioning	(Rozado 2020) (Papakyriakopoulos et al. 2020)
	Race	(Obermeyer et al. 2019) (Martínez-Plumed et al. 2021) (Schuler and Montardo 2020) (Khalil et al. 2020) (Ostherr 2020) (Gupta et al. 2021) (Zajko 2021) (Mcquillan 2018) (Evangelo and Oliveira 2021) (Bonchi et al. 2017) (Rozado 2020)
	Religion	(Rozado 2020) (Akter et al. 2021)
	Intersecctional	(Xenidis 2020)

Table 1. Thematic categories from the titles and abstracts. Source: authors, 2023 (continued on next page).

344   Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

Ambiences and contexts	Healthcare systems (diagnosis, terapeutics, prognosis)	(Obermeyer et al. 2019) (kerasidou 2021) (Starke et al. 2021) (Norori et al. 2021) (Paulus and Kent 2020) (Walsh et al. 2020)
	Criminal systems	(Castro 2020)
	Corporate systems	(Bolander 2019) (Gangadharan and Niklas 2019)
	Electoral systems	(Unkel and Haim 2019)
	Legal systems	(Gutiérrez 2021) (Xenidis 2020) (Rhoen and Feng 2018)
	Social media (Twitter, Facebook, Instagram, Youtube)	(Fosch-Villaronga et al. 2021), (Rettberg 2020) (Costello et al. 2021) (Are 2021) (Abul-Fottouh et al. 2020)
	Search engines (Yandex, Google)	(Kravets and Toepfl 2021) (Zhang et al. 2019)
	Facial recognition systems	(Khalil et al. 2020)
	Biometrics / Emotion AI	(Mantello et al. 2021)
	UM Sustainable Development Goals	(Ryan et al. 2020)
Types of algorithms	Collaborative filtering	(McQuillan 2018) (Nader 2020)
	Social rankings	(Evangelo and Oliveira 2021) (Zhang et al. 2021)
	Social media content recommendation	(Abul-Fottouh et al. 2020) (Papakyriakopoulos et al. 2020) (Gupta et al. 2021)
	Public Relevance	(Schuler and Montardo 2020)
	Predictive	(Paulus and Kent 2020) (Fennell et al. 2019) (Xenidis 2020)
	Randomized classifyers	(Oneto et al. 2020)
	Discrimination Prevention and Rule Protection Algoritms	(Babu and Pushpa 2018)
Types of data	Personal data	(Rettberg 2020)
	Behavioral data	(Fennell et al. 2019) (Costello et al. 2021)
	Sensitive data	(Rhoen and Feng 2018) (Oneto et al. 2020)
Specific Tools and instruments	DiscriLens	(Wang et al. 2021)
	Outside in: exile at home	(Castro 2020)
	SensitiveNets	(Morales 2019)
	Situated Data Analysis	(Rettberg 2020)
Methodologies / Procedures	Feature Space Decomposition – FSD	(Fennell et al. 2019)
Interactions and consequences	Opinion formation	(Peralta et al. 2021)
	Algorithmic eco-systems	(Papakyriakopoulos 2022)

Table 1 (continued from previous page)

Knowl. Org. 50(2023)No.5

345

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

In this sense, it is observed that the ethical values of Freedom, Information safety and Privacy aim to safeguard the subject of personal data. In turn, the values of Cultural warrant, Equity, Reliability, and Risk minimization must be expressed in the configuration of the algorithms. Regarding ethical problems, Censorship, Defamation, Digital segregation, Marginalization, Prejudices, Racism, Surveillance and Violence refer to the possibility of damages to be suffered by the subject of personal data while the ethical problems of Idiosyncrasy, Lack of cultural warrant, and Misrepresentation should be avoided when planning algorithm configuration.

5.0 Conclusion

The analyzed knowledge domain, related to personal data in its behavioral dimension and incorporating sensitive data, proved to be susceptible to the action of relevance, filtering, predictive, social ranking, content recommendation and random classification algorithms. Such algorithms, in turn, may have built-in, in their programming, discriminatory biases related to gender, sexual orientation, race, nationality, religion, age, social class, socioeconomic profile, physical appearance, and political position.

Approaches applied from the literature show the prevalence of concerns in specific environments such as Health Systems, Legal Systems, Organizational Systems, Electoral Systems, Social Media, and Biometric Systems.

The content analysis carried out on the corpus allowed thematic categorization (categories and subcategories) of the approaches given to the topic in the analyzed literature. These categories, in turn, contribute to the conceptual delimitation of the analyzed topic, as they highlight the dimensions in which it is scientifically treated, revealing a multifaceted nature that incorporates elements of an axiological, contextual, technological, instrumental, methodological nature, and, consequently, interdisciplinary.

This multifaceted nature, in turn, highlights how the issue of personal data protection is included in the research agenda of different areas of Information Science, affecting not only the organization and representation of knowledge, as discussed in this work, but also with consequences in fields such as information literacy, information retrieval, data descriptive processes, and even the formation and development of collections in information units, among others.

Regarding the ethical consequences of this theme, values related to transparency, responsibility, justice, reliability and algorithmic interpretability are evident, as well as a concern with problems related to algorithmic colonization, oppressive digital practices and, ultimately, the so-called "surveillance capitalism". As a result, this context can lead to consequences relating to the effects of algorithms on information transfer and opinion formation, as well as the degree of interaction between people and algorithms (in so-called socio-algorithmic ecosystems) and the development of policy in social structures.

That said, it becomes evident that the algorithms using personal data as a knowledge base, organize and name them to generate analyses and decision-making susceptible to biases of various orders – arising from their programming process – which might culminate in discriminatory situations for a particular subject or group, as they do not take into account primordial issues related to social equity in the construction of the rules that govern the algorithm. This directly affects the universe of Information Science in aspects such as data lifecycle management, development of the user's critical capacity and, more especially, knowledge organization. Such aspects lead to the need for more vertical studies on specific topics such as information self-determination, categorization and comparison of biases that arise from algorithmic activities and the role of knowledge organization tools to support the construction of personal databases.

The original work by Guimarães et al. (2008) regarding the identification and categorization of values and ethical problems in organization and representation of knowledge was published prior to the promulgation of the GDPR, at a time when protection of personal data was not yet an object of the area's research agenda; since then, an expansion/folding of the area's ethical universe has been observed. In this sense, if, on the one hand, values such as Privacy and Autonomy, as well as problems such as Surveillance, Digital segregation, Racism, and Marginalization remain in this new reality, new ethical values emerge, such as Algorithmic Responsibility, Algorithmic Transparency, Justice and Algorithmic Trust, as well as new ethical problems such as Algorithmic Colonization, Oppressive Digital Practices, Facial Recognition Systems, and Social Rankings, which makes the axiological universe of this field even more complex.

Endnotes

1. The helical movement has a three-dimensional nature differently from a traditional spiral movement with a two-dimensional nature.
2. These data need to be periodically updated in order to be always complete, correct and real.
3. Nevertheless, only the technical knowledge is not enough to motivate social changes, citizens need also be able to questioning their cultures and social conditions (Masur, 2020).
4. This method is composed by 3 steps: pre-analysis (composition of the corpus and formulating hypotheses), exploration (reading, extraction and description of contents), and treatment of results (application of thematic categories to the extracted content) (Guimarães; Sales, 2010).

346   Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

5. Orestis Papakyriakopoulos "studied civil engineering at the Technical University of Athens, philosophy of science and technology at the Technical University of Munich and defended my PhD in computer science at the same institution". He is a "researcher in Political Data Science, analyzing new and old media by the application of data-intensive algorithms, as well as the political impact of the use of data-intensive algorithms themselves" at the Center for Information Technology Policy at Princeton University (https://civicmachines. com/about.html)

6. Authors' countries: USA (49); UK (17); Spain (15); Switzerland (12); China (8); Germany, Hong Kong, Indonesia and Netherlands (6 each); Australia and India (5 each); Brazil, Canada, Ireland, Italy and Arabic Emirates (4 each); Norway and Vietnam (3 each); Austria, Denmark, Japan, Mexico and South Korea (2 each); and New Zealand, Chile, Cyprus, Finland, Russia, South Africa and Sweden (1 each).

7. World Division according to socio-economic characteristics. Most of these countries are located in the Southern Hemisphere (Litonjua, 2012).

8. For that, two conceptual domains of keywords were identified: Ethics (represented by the terms Ethics; Ethical; Ethos; Deontology; Value(s); Conduct; Moral) and Knowledge Organization (represented by the terms Knowledge Organization; Knowledge Representation; Indexing; Classification; Subject Cataloguing; Subject Analysis; Content Analysis; Abstracting; Thesaurus(i) and Subject Headings).

9. Ethical values: Accessibility, Autonomy, Authorship (copyright), Competence skill, Consistency, Co-operation, Cultural warrant, Efficiency, Equity, Exhaustivity, Flexibility, Freedom, Hospitality, Information diversity, Information safety, Power awareness, Precision, Privacy, Professional recognition, Reliability, Risk minimization, Up-to-dating, and Usability.

10. Ethical problems: Ambiguity, Biased translations, Censorship, Defamation, Digital segregation, Discharge of electronic garbage, Idiosyncrasy, Impartiality or neutrality belief, Inaccessibility to information, Informational Directness, Lack of cultural warrant, Marginalization, Misrepresentation, Negligence, Pornography, Prejudices in terminology, Professional inefficiency, Professionals being replaced by technology, Racism, Surveillance, and Violence.

Acknowledgements

References

Abul-Fottouh, Deena, Melodie Yunju Song, and Anatoliy Gruzd. 2020. "Examining Algorithmic Biases in YouTube's Recommendations of Vaccine Videos." *International Journal of Medical Informatics* 140 (August): 104175. https://doi.org/10.1016/j.ijmedinf.2020.104175.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature* 54, no. 2: 442–92. https://doi.org/10.1257/jel.54.2.442.

Akter, Shahriar, Yogesh K. Dwivedi, Kumar Biswas, Katina Michael, Ruwan J. Bandara, and Shahriar Sajib. 2021. "Addressing Algorithmic Bias in AI-Driven Customer Management." *Journal of Global Information Management* (JGIM) 29, no. 6: 1–27. https://doi.org/10.4018/JGIM.20211101.oa3.

Are, Carolina. 2021. "The Shadowban Cycle: An Autoethnography of Pole Dancing, Nudity and Censorship on Instagram." *Feminist Media Studies* 0, no.0: 1–18. https://doi.org/10.1080/14680777.2021.1928259.

Babu, Mylam, and Sankaralingam Pushpa. 2018. "An Efficient Discrimination Prevention and Rule Protection Algorithms Avoid Direct and Indirect Data Discrimination in Web Mining." *International Journal of Intelligent Engineering and Systems* 11, no.4: 212–20. https://doi.org/10.22266/ijies2018.0831.21.

Bagatini, José Augusto, José Augusto Chaves Guimarães, and Ricardo César Gonçalves Sant'Ana. 2021. "Gerenciamento dos Dados Pessoais em Arquivos: Uma Perspectiva Centrada no Indivíduo Com Base na LGPD". [Archival Management of Personal Data: a Human-Centred Perspective Based on GDPR]" *Acervo* 34, no.3: 1–20. https://revista.an.gov.br/index.php/revistaacervo/article/view/1749/1679

Bardin, Laurence. 2013. *L'analyse de Contenu.* Paris: Presses Universitaires de France.

Bartsch, Miriam, and Tobias Dienlin. 2016. "Control Your Facebook: An Analysis of Online Privacy Literacy." *Computers in Human Behavior* 56: 147–54. https://doi.org/10.1016/j.chb.2015.11.022.

Bauman, Zygmunt, and David Lyon. 2012. *Liquid Surveillance: A Conversation.* Cambridge: Polity.

Beghtol, Clare. 2002. "A Proposed Ethical Warrant for Global Knowledge Representation and Organization Systems." *Journal of Documentation* 58, no.5: 507–32. https://doi.org/10.1108/00220410210441

Beghtol, Clare. 2005. "Ethical Decision-Making for Knowledge Representation and Organization Systems for Global Use." *Journal of the American Society for Information Science and Technology* 56, no.9: 903–12. https://doi.org/10.1002/asi.20184

Knowl. Org. 50(2023)No.5

347

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

Begley, Keith, Cecily Begley, and Valerie Smith. 2021. "Shared Decision-Making and Maternity Care in the Deep Learning Age: Acknowledging and Overcoming Inherited Defeaters." *Journal of Evaluation in Clinical Practice* 27, no.3: 497–503. https://doi.org/10.1111/jep.13515.

Berman, Sanford. 1993. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. 2nd ed. Jefferson; North Carolina: London: McFarland & Company Inc. Publishers.

Black, Edwin. 2001. *IBM and the Holocaust: The Strategic Alliance between Nazi Germany and America's Most Powerful Corporation*. Amsterdam: Elsevier.

Bolander, Thomas. 2019. "What Do We Loose When Machines Take the Decisions?" *Journal of Management and Governance* 23, no. 4: 849–67. https://doi.org/10.1007/s10997-019-09493-x.

Bonchi, Francesco, Bud Mishra, and Daniele Ramazzotti. 2017. "Exposing the Probabilistic Causal Structure of Discrimination." *International Journal of Data Science and Analytics* 33: 1–21. https://doi.org/10.1007/s41060-016-0040-z

Castells, Manuel. 2009. *The Rise of the Network Society*. 2nd ed. Oxford: Wiley-blackwell.

Castro, Annabel. 2020. "Outside in: Exile at Home an Algorithmic Discrimination System." *Artnodes* no.26: 1-7. https://doi.org/10.7238/a.v0i26.3359

Cavoukian, Ann. 2006. "Privacy by Design: The 7 Foundational Principles." Information & Privacy Commissioner of Ontario. https://iapp.org/media/pdf/resource_center/pbd_implement_7found_principles.pdf.

Costello, Cory, Sanjay Srivastava, Reza Rejaie, and Maureen Zalewski. 2021. "Predicting Mental Health From Followed Accounts on Twitter." *Collabra*: *Psychology* 7, no.1: 18731. https://doi.org/10.1525/collabra.18731.

Du, Mengnan, Fan Yang, Na Zou, and Xia Hu. 2020. "Fairness in Deep Learning: A Computational Perspective." arXiv:1908.08843 [Cs, Stat], March. http://arxiv.org/abs/1908.08843.

Evangelo, Naiara Silva, and Fátima Cristina Regis Martins de Oliveira. 2021. "The Black Social Ranking Experience at Uber: a Racialized Reflection on Contemporary Surveillance." *Comunicação e Sociedade* 39 (June): 83–100. https://doi.org/10.17231/comsoc.39(2021).2838.

Federal Trade Commission. 2014. "Data Brokers: A Call for Transparency and Accountability." Washington. https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014.

Fennell, Peter G., Zhiya Zuo, and Kristina Lerman. 2019. "Predicting and Explaining Behavioral Data with Structured Feature Space Decomposition." *EPJ Data Science* 8, no.1: 23. https://doi.org/10.1140/epjds/s13688-019-0201-0.

Fernández-Molina, Juan-Carlos, José Guimarães, Silvana Vidotti, Adriana Flamino, Alexandre Souza, Liriane Camargo, Marcel Silva, Patricia Moreno, and Rogério Ramalho. 2005. "Aspectos Éticos de Las Nuevas Tecnologías de La Información y La Comunicación: Y Su Reflejo En La Organización y Representación Del Conocimiento". [Ethical Aspects of The New Information Technologies and Their Reflexes to Knowledge Organizaiton and Representation] In *La Dimensión Humana de La Organización Del Conocimiento*, edited by Jesús Gascón, Ferran Burguillos, and Amadeu Pons. Barcelona: Universitat de Barcelona, 177–86

Fosch-Villaronga, E., A. Poulsen, R. A. Søraa, and B. H. M. Custers. 2021. "A Little Bird Told Me Your Gender: Gender Inferences in Social Media." *Information Processing & Management* 58, no.3: 102541. https://doi.org/10.1016/j.ipm.2021.102541.

Gangadharan, Seeta Peña, and Jędrzej Niklas. 2019. "Decentering Technology in Discourse on Discrimination." *Information, Communication & Society* 22, no.7: 882–99. https://doi.org/10.1080/1369118X.2019.1593484.

Gardner, Carolyn Caffey. 2019. "Teaching Algorithmic Bias in a Credit-Bearing Course." *International Information & Library Review* 51, no.4: 321–27. https://doi.org/10.1080/10572317.2019.1669937.

Greenwald, Glenn, and Ewen MacAskill. 2013. "NSA Prism Program Taps in to User Data of Apple, Google and Others," *The Guardian*, June 7, 2013, sec. US news. https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data.

Guimarães, José Augusto Chaves. 2008. "A Dimensão Teórica do Tratamento Temático da Informação e suas interlocuções com o Universo Científico da International Society for Knowledge Organization (ISKO)." [The Theoretical Dimension of Subject Treatment of Infroamtion and its Dialogues with the Scientific Universe of ISKO] *Revista Ibero-Americana de Ciência Da Informação* 1, no.1: 77–9. https://doi.org/10.26512/rici.v1.n1.2008.940

Guimarães, José Augusto Chaves. 2014. "Análise de Domínio como Perspectiva Metodológica em Organização da Informação" [Domain Analysis as a Methodological Perspective in Information Organization]. *Ciência Da Informação* 41, no.1: 13–21. https://doi.org/10.18225/ci.inf.v43i1.1415

Guimarães, José Augusto Chaves. 2017. "Slanted Knowledge Organization as a New Ethical Perspective." In *The Organization of Knowledge, Studies in Information*, volume 12 edited by Jack Andersen and Laura Skouvig. Emerald Publishing Limited, 87–102.https://doi.org/10.1108/S2055-537720170000011012.

Guimarães, José Augusto Chaves, Juan Carlos Fernández-Molina, Fabio Assis Pinho, and Suellen de Oliveira

348

Knowl. Org. 50(2023)No.5
J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

Milani. 2008. "Ethics in the Knowledge Organization Environment: An Overview of Values and Problems in the LIS Literature." In *Cultural and Identity in Knowledge Organization: Proceedings of the Tenth International Isko Conference 5-8 August, 2008 Montréal Canada*, edited by Clément Arsenault and Joseph T. Tennis. Advances in knowledge organization 11. Würzburg: Ergon Verlag,340–46.

Guimarães, José Augusto Chaves, and Rodrigo de Sales. 2010. "Análise documental: concepções do universo acadêmico brasileiro em Ciência da Informação". [Conceptions of Document Analysis in the Brazilin IS Environment]. *Datagramazero*. http://www.dgz.org.br/fev10/Art_02.htm.

Gupta, Manjul, Carlos M. Parra, and Denis Dennehy. 2021. "Questioning Racial and Gender Bias in AI-Based Recommendations: Do Espoused National Cultural Values Matter?" *Information Systems Frontiers*, June. https://doi.org/10.1007/s10796-021-10156-2.

Gutiérrez, Antonio García. 2002. "Knowledge Organization from a 'Culture of the Border': Towards a Transcultural Ethics of Mediation." In *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries: Proceedings of the seventh International Isko Conference 10-13 July Granada Spain*, edited by María José López-Huertas. Advances in Knowledge Organization 8. Würzburg: Ergon Verlag, 516–22.

Gutiérrez, Catherine Muñoz. 2021. "La discriminación en una sociedad automatizada: Contribuciones desde América Latina." *Revista Chilena de Derecho y Tecnología* 10, no.1: 271–307. https://doi.org/10.5354/0719-2584.2021.58793.

Hjørland, Birger. 2002. "Domain Analysis in Information Science. Eleven Approaches — Traditional as Well as Innovative." *Journal of Documentation* 58, no.4: 422–62. https://doi.org/ 10.1108/00220410210431136

Hjørland, Birger. 2017. "Domain Analysis." *Knowledge Organization* 44, no.6: 436–64. https://doi.org/10.5771/0943-7444-2017-6-436

Hjørland, Birger, and Hanne Albrechtsen. 1995. "Toward a New Horizon in Information Science: Domain Analysis." *Journal of the American Society for Information Science* 46, no.6: 400–25. https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y

Instituto Humanitas Unisinos. 2016. "Sistema de Algoritmo que Determina Pena de Condenados Cria Polêmica Nos EUA." 2016. http://www.ihu.unisinos.br/78-noticias/561869-sistema-de-algoritmo-que-determina-pena-de-condenados-cria-polemica-nos-eua.

Jooste, Yvonne. 2021. "Surveillance Capitalism as White World-Making." *Acta Academica* 53, no.1: 44–67. https://doi.org/10.18820/24150479/aa53i1.3

Kehrenberg, T., Zexun Chen, and Novi Quadrianto. 2020. "Tuning Fairness by Balancing Target Labels". *Artificial Intelligence* 12: https://doi.org/10.3389/frai.2020.00033.

Kerasidou, Angeliki. 2021. "Ethics of Artificial Intelligence in Global Health: Explainability, Algorithmic Bias and Trust*." Journal of Oral Biology and Craniofacial Research* 11, no.4: 612–14. https://doi.org/10.1016/j.jobcr.2021.09.004.

Ketelaar, Eric. 2014. "Archives, Memories and Identities." In *Archives and Recordkeeping: Theory into Practice*, edited by Caroline Brown. London: Facet Publishing.

Khalil, Ashraf, Soha Glal Ahmed, Asad Masood Khattak, and Nabeel Al-Qirim. 2020. "Investigating Bias in Facial Analysis Systems: A Systematic Review." *IEEE Access* 8: 130751–61. https://doi.org/10.1109/ACCESS.2020.3006051.

Kravets, Daria, and F. Toepfl. 2021. "Gauging Reference and Source Bias over Time: How Russia's Partially State-Controlled Search Engine Yandex Mediated an Anti-Regime Protest Event." *Information, Communication & Society* 0, no.0: 1–17. https://doi.org/10.1080/1369118X.2021.1933563.

Kristiadi, Dedy Prasetya, Po Abas Sunarya, Melvin Ismanto, Joshua Dylan, Ignasius Raffael Santoso, and Harco Leslie Hendric Spits Warnars. 2020. "Framework for Developing Algorithmic Fairness." *Bulletin of Electrical Engineering and Informatics* 9, no.4: 1550–57. https://doi.org/10.11591/eei.v9i4.2028.

Langenderfer, Jeff, And Anthony D. Miyazaki. 2009. "Privacy in the Information Economy." *The Journal of Consumer Affairs* 43, no. 3: 380–88.

Litonjua, M. D. 2012. "Third World/Global South: From Modernization, to Dependency/Liberation, to Postdevelopment." *Journal of Third World Studies; Americus* 29, no.1: 25–56.

Mantello, Peter, Manh-Tung Ho, Minh-Hoang Nguyen, and Quan-Hoang Vuong. 2021. "Bosses without a Heart: Socio-Demographic and Cross-Cultural Determinants of Attitude toward Emotional AI in the Workplace." *AI & Society*, November. https://doi.org/10.1007/s00146-021-01290-1.

Martínez-Plumed, F., C. Ferri, D. Nieves, and J. Hernández-Orallo. 2021. "Missing the Missing Values: The Ugly Duckling of Fairness in Machine Learning." *International Journal of Intelligent Systems* 36, no.7: 3217–58. https://doi.org/10.1002/int.22415.

Masur, Philipp K. 2020. "How Online Privacy Literacy Supports Self-Data Protection and Self-Determination in the Age of Information." *Media and Communication* 8, no.2: 258–69. https://doi.org/10.17645/mac.v8i2.2855.

Knowl. Org. 50(2023)No.5

349

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

Mayer-Schönberger, Viktor. 2011. "Delete: The Virtue of Forgetting in the Digital Age." In *The Demise of Forgetting – and Its Drivers*, 4th ed. Princeton: Princeton University Press.

McQuillan, Dan. 2018. "Data Science as Machinic Neoplatonism." *Philosophy & Technology* 31, no.2: 253–72. https://doi.org/10.1007/s13347-017-0273-3.

Menárguez, Ana Torres. 2018. "Os Privilegiados São Analisados Por Pessoas; As Massas, Por Máquinas," *El País*. November 21, 2018. https://brasil.elpais.com/brasil/2018/11/12/tecnologia/1542018368_035000.html.

Milani, Suellen Oliveira, and José Augusto Chaves Guimarães. 2017. "Problemas Relacionados a Biases Em Sistemas de Organização do Conhecimento: Perspectivas para a Representação de Assunto" [Biases in KOS: Perspectives for Subject Representation]. In *Memória, Tecnologia e Cultura Na Organização Do Conhecimento*, edited by Fabio Assis Pinho and José Augusto Chaves Guimarães. Estudos Avançados em Organização do Conhecimento 4. Recife: Ed. UFPE, 399–407.

Morales, Débora. 2019. "Viés Algorítmico Na Interpretação de Dados." *Campo Grande News* (blog). 2019. https://www.campograndenews.com.br/artigos/vies-algoritmico-na-interpretacao-de-dados.

Nader, Karim. 2020. "Dating through the Filters." *Social Philosophy and Policy* 37, no.2: 237–48. https://doi.org/10.1017/S0265052521000133.

Norori, Natalia, Qiyang Hu, Florence Marcelle Aellen, Francesca Dalia Faraci, and Athina Tzovara. 2021. "Addressing Bias in Big Data and AI for Health Care: A Call for Open Science." *Patterns* 2, no.10: 100347. https://doi.org/10.1016/j.patter.2021.100347.

Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366 (6464): 447–53. https://doi.org/10.1126/science.aax2342.

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht: Kluwer.

O'Neil, Cath. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Oneto, Luca, Michele Donini, Massimiliano Pontil, and John Shawe-Taylor. 2020. "Randomized Learning and Generalization of Fair and Private Classifiers: From PAC-Bayes to Stability and Differential Privacy." *Neurocomputing* 416 (November): 231–43. https://doi.org/10.1016/j.neucom.2019.12.137.

Ostherr, Kirsten. 2020. "Artificial Intelligence and Medical Humanities." *Journal of Medical Humanities*, July. https://doi.org/10.1007/s10912-020-09636-4.

Papakyriakopoulos, Orestis. 2022. "Political Machines: A Framework for Studying Politics in Social Machines." *AI & Society* 37, no.1: 113–30. https://doi.org/10.1007/s00146-021-01180-6.

Papakyriakopoulos, Orestis, Juan Carlos Medina Serrano, and Simon Hegelich. 2020. "Political Communication on Social Media: A Tale of Hyperactive Users and Bias in Recommender Systems." *Online Social Networks and Media* 15 (January): 100058. https://doi.org/10.1016/j.osnem.2019.100058.

Paulus, Jessica K., and David M. Kent. 2020. "Predictably Unequal: Understanding and Addressing Concerns That Algorithmic Clinical Prediction May Increase Health Disparities." *Npj Digital Medicine* 3, no.1: 1–8. https://doi.org/10.1038/s41746-020-0304-9.

Peralta, A. F., M. Neri, J. Kertész, and G. Iñiguez. 2021. "Effect of Algorithmic Bias and Network Structure on Coexistence, Consensus, and Polarization of Opinions." *Physical Review E* 104 (4). https://journals.aps.org/pre/abstract/10.1103/PhysRevE.104.044312.

Poikola, Antti, Kai Kuikkaniemi, and Harri Honko. 2020. "MyData – Um Modelo Nórdico Para Gestão e Processamento de Dados Pessoais Centrado No Ser Humano". Rio de Janeiro. https://mydatabrasil.wordpress.com/2020/09/01/mydata-um-modelo-nordico-para-gestao-e-processamento-de-dados-pessoais-centrado-no-ser-humano/.

Pombo, Olga. 1998. "Da Classificação dos Seres à Classificação dos Saberes." *Leituras, Revista Da Biblioteca Nacional de Lisboa* 2: 19–33.

Rettberg, Jill Walker. 2020. "Situated Data Analysis: A New Method for Analysing Encoded Power Relationships in Social Media Platforms and Apps." *Humanities and Social Sciences Communications* 7. no.1: 1–13. https://doi.org/10.1057/s41599-020-0495-3.

Rhoen, Michiel, and Qing Yi Feng. 2018. "Why the 'Computer Says No': Illustrating Big Data's Discrimination Risk through Complex Systems Science." *International Data Privacy Law* 8, no.2: 140–59. https://doi.org/10.1093/idpl/ipy005.

Romansky, Radi. 2015. "Social Computing and Digital Privacy." *Communication & Cognition* 48, no. 34: 65–82.

Rosenbaum, Howard, and Pnina Fichman. 2019. "Algorithmic Accountability and Digital Justice: A Critical Assessment of Technical and Sociotechnical Approaches." *Proceedings of the Association for Information Science and Technology* 56, no.1: 237–44. https://doi.org/10.1002/pra2.19.

Rozado, David. 2020. "Wide Range Screening of Algorithmic Bias in Word Embedding Models Using Large Sentiment Lexicons Reveals Underreported Bias Types." *Plos One* 15 (4): e0231189. https://doi.org/10.1371/journal.pone.0231189.

Ryan, Mark, Josephina Antoniou, Laurence Brooks, Tilimbe Jiya, Kevin Macnish, and Bernd Stahl. 2020.

350

Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

"The Ethical Balance of Using Smart Information Systems for Promoting the United Nations' Sustainable Development Goals." *Sustainability* 12, no.12: 4826. https://doi.org/10.3390/su12124826.

Sales, Rodrigo de. 2008. "Tesauros e Ontologias sob a luz da Teoria Comunicativa da Terminologia" [Thesauri and Ontologies under the Perspective of the Communicative Theory of Terminology]." PhD diss., Universidade Federal de Santa Catarina, 2008. https://repositorio.ufsc.br/bitstream/handle/123456789/92173/251391.pdf?sequence=1&isAllowed=y

San Segundo, Rosa. 1996. "Sistemas de Organización Del Conocimiento." Madrid: *Bolentin Oficial Del Estado*; Universidad Carlos III de Madrid.

Sant'Ana, Ricardo César Gonçalves. 2016. "Ciclo de Vida dos Dados: Uma Perspectiva a Partir da Ciência da Informação". [The Data Life Cycle under the Perspective of Information Science] *Informação & Informação* 21, no.2: 116-42. https://doi.org/10.5433/1981-8920.2016v21n2p116.

Saulles, Martin de. 2015. *Information 2.0: New Models of Information Production, Distribution and Consuption*. 2nd ed. London: Facet.

Schuler, Luis Gustavo Bier, and Sandra Portella Montardo. 2020. "A Política das Máquinas: Vieses em Algoritmos de Relevância Pública." *Revista Tecnologia e Sociedade* 16, no.45: 300–12. https://doi.org/10.3895/rts.v16n45.10380.

Shen, Hong, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. "Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors." *Proceedings of the ACM on Human-Computer Interaction 5 (CSCW2)*: 1–29. https://doi.org/10.1145/3479577.

Smit, Johanna Wilhelmina, and Aldo de Albuquerque Barreto. 2002. "A Ciência Da Informação: Base Conceitual Para a Formação do Profissional" [Information Science: Conceptual Bases for Professional Education]. In *Formação do Profissional da Informação*, edited by Marta Ligia Pomim Valentim. São Paulo: Polis, 9-23.

Starke, Georg, Eva De Clercq, and Bernice S. Elger. 2021. "Towards a Pragmatist Dealing with Algorithmic Bias in Medical Machine Learning." *Medicine, Health Care and Philosophy* 24, no.3: 341–49. https://doi.org/10.1007/s11019-021-10008-5.

Sun, W., O. Nasraoui, and P. Shafto. 2020. "Evolution and Impact of Bias in Human and Machine Learning Algorithm Interaction." *PlosOne*. https://doi.org/10.1371/journal.pone.0235502.

Tennis, Joseph T. 2003. "Two Axes of Domains for Domain Analysis." *Knowledge Organization* 30, nos.3/4: 191–95. https://doi.org/10.5771/0943-7444-2003-3-4-191

UNCTAD. 2020. "Data Protection and Privacy Legislation Worldwide | UNCTAD." 2020. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide.

Unkel, J., and M. Haim. 2019. "Googling Politics: Parties, Sources, and Issue Ownerships on Google in the 2017 German Federal Election Campaign – Julian Unkel, Mario Haim, 2021." *Social Science Computer Review* 39, no. 5:844-61. https://doi.org/10.1177/0894439319881634.

Van Dijck, José. 2007. *Mediated Memories in the Digital Age*. Stanford: Stanford University Press.

Véliz, Carissa. 2021. *Privacy Is Power: Why and How You Should Take Back Control of Your Data*. 1st ed. London: Bantam Press.

Verma, Sonia, and Sanghmitra Sheel Acharya. 2018. "Social Identity and Perceptions about Health Care Service Provisioning by and for the Dalits in India." *Social Identities* 24, no.3: 327–38. https://doi.org/10.1080/13504630.2017.1376280.

Walsh, C. G., B. Chaudhry, P. Dua, K. W. Goodman, B. Kaplan, R. Kavuluru, A. Solomonides, and V. Subbian. 2020. "Stigma, Biomarkers, and Algorithmic Bias: Recommendations for Precision Behavioral Health with Artificial Intelligence." *Jamia Open* 3, no.1: 9–15. https://doi.org/10.1093/jamiaopen/ooz054.

Wang, Qianwen, Zhenhua Xu, Zhutian Chen, Yong Wang, Shixia Liu, and Huamin Qu. 2021. "Visual Analysis of Discrimination in Machine Learning." *IEEE Transactions on Visualization and Computer Graphics* 27, no.2: 1470–80. https://doi.org/10.1109/TVCG.2020.3030471.

Westin, Alan. 1967. *Privacy and Freedom*. Nova York: Ig Publishing.

Williams, Betsy Anne, Catherine F. Brooks, and Yotam Shmargad. 2018. "How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications." *Journal of Information Policy* 8: 78–115. https://doi.org/10.5325/jinfopoli.8.2018.0078.

Wissinger, Christina L. 2017. "Privacy Literacy: From Theory to Practice." *Communications in Information Literacy* 11, no.2: 378–89.

World Economic Forum. 2021. "The EDISON Alliance." 2021. https://www.weforum.org/the-edison-alliance/.

Xenidis, Raphaële. 2020. "Tuning EU Equality Law to Algorithmic Discrimination: Three Pathways to Resilience." *Maastricht Journal of European and Comparative Law* 27, no.6: 736–58. https://doi.org/10.1177/1023263X20982173.

Yan, Spring. 2021. "Algorithms Are Not Bias-Free: Four Mini-Cases." *Human Behavior and Emerging Technologies* 3, no.5: 1180–84. https://doi.org/10.1002/hbe2.289.

Knowl. Org. 50(2023)No.5

J. A. Bagatini, J. Augusto Chaves Guimarães. Algorithmic Discriminations and Their Ethical Impacts on Knowledge Organization

351

Yu, Shi-Eun, Jin-Sup Eom, and Woo-Taek Jeon. 2012. "The Factors Affecting the Development of National Identity as South Korean in North Korean Refugees Living in South Korea." T201203155.pdf 9 (3): 209–16.

Zajko, Mike. 2021. "Conservative AI and Social Inequality: Conceptualizing Alternatives to Bias through Social Theory." *AI & Society* 36. no.3: 1047–56. https://doi.org/10.1007/s00146-021-01153-9.

Zhang, Shilun, Matúš Medo, Linyuan Lü, and Manuel Sebastian Mariani. 2019. "The Long-Term Impact of Ranking Algorithms in Growing Networks." *Information Sciences* 488 (July): 257–71. https://doi.org/10.1016/j.ins.2019.03.021.

Zhang, Yu, Peter Tiňo, Aleš Leonardis, and Ke Tang. 2021. "A Survey on Neural Network Interpretability." *IEEE Transactions on Emerging Topics in Computational Intelligence* 5, no.5: 726–42. https://doi.org/10.1109/TETCI.2021.3100641.