

ALGORITHMIC BIAS AND INTELLECTUAL FREEDOM: NAVIGATING THE ETHICAL CHALLENGES OF ARTIFICIAL INTELLIGENCE -DRIVEN CURATION IN DIGITAL LIBRARIES

¹Dr. Ezema, Anezi Leonard, CLN, ²Dr. OGBUABOR, Chimezie Damian, CLN, &

³Dr. Uzomba, Christian Emeka, CLN

¹Cataloguing and Classification Department, MOUAAU

²Department of library & information Science UNN

³Federal University Oyo-Ekiti

¹leonardanezi@gmail.com ²Damian.ogbuabor@unn.edu.ng ³emeka.uzomba@fuoye.edu.ng

Abstract

The integration of artificial intelligence (AI) into digital libraries has transformed how information is organized, accessed, and recommended. Artificial intelligence (AI)-driven curation tools, using machine learning and natural language processing, enable highly personalized experiences, streamline retrieval, and manage vast collections efficiently. However, alongside these benefits arise critical ethical challenges, particularly concerns about algorithmic bias and threats to intellectual freedom issues central to the mission of libraries. This paper critically examines how biases in AI systems, whether stemming from training data, design decisions, or user interactions, can result in inequitable information access, reinforce societal inequalities, and narrow the range of perspectives available to users. As digital libraries increasingly adopt AI, it is vital to ensure that technological advancements support, rather than undermine, democratic values. The discussion highlights strategies for mitigating ethical risks, such as promoting ethical design principles, ensuring transparency in algorithmic processes, using inclusive and representative datasets, and strengthening institutional oversight. Upholding intellectual freedom in the era of AI requires libraries to navigate these challenges thoughtfully, maintaining their role as defenders of equitable and diverse access to information in an increasingly complex digital landscape.

Keywords: Algorithmic bias, intellectual freedom, digital libraries, ethics challenges, Artificial Intelligence –driven, Curation.

Introduction

The incorporation of artificial intelligence (AI) into digital library environments has led to a profound evolution in how information is curated, managed, and disseminated. Historically, librarians played a central role in tasks like cataloguing, creating metadata, retrieving information, and curating collections. However, with the advent of sophisticated machine

learning (ML) techniques, many of these functions are now performed faster and at a much larger scale (Murphy, Khosrow-Pour, & Khosrow-Pour, 2021). Tools such as natural language processing (NLP), predictive modeling, and recommendation systems have significantly reshaped the ways users interact with digital libraries, offering more personalised and efficient access to resources. Through machine learning, libraries can forecast user needs, suggest relevant materials, automate metadata production, and detect emerging research trends, enabling a broader and more efficient response to the diverse needs of users (Beel, Gipp, Langer, & Breitinger, 2016).

Beyond operational improvements, AI is also facilitating the democratization of information access by bridging linguistic, geographic, and disciplinary divides (Gonzalez-Perez, Flores-Vivar, & Lopez-Golán, 2020). Nevertheless, the deployment of AI in digital libraries introduced significant ethical challenges, particularly related to algorithmic bias and the preservation of intellectual freedom. Algorithmic bias, characterised by systemic inaccuracies that resulted in unfair outcomes, often arises from historical prejudices embedded in training datasets or design shortcomings (Mehrabi, Morstatter, Saxena, Lerman, & Galstyan, 2021). In the context of digital libraries, such biases can lead to the diminished visibility of minority perspectives, the sidelining of controversial topics, and discrimination across racial, gender, or regional lines, ultimately threatening libraries' mission to provide equitable access to diverse knowledge.

The issue of safeguarding intellectual freedom, a foundational tenet of librarianship, is equally at stake. Intellectual freedom supports the individual's right to seek and receive information without censorship or bias (American Library Association, 2017). Libraries have traditionally protected a diversity of viewpoints within their collections. However, AI particularly through recommendation algorithms may unintentionally narrow the range of perspectives offered. By prioritizing materials that align with popular or mainstream preferences, these systems can create "filter bubbles," limiting users to content that reaffirms existing beliefs and restricting opportunities for critical engagement (Noble, 2018). Such narrowing ultimately hinders scholarly inquiry, critical thinking, and democratic participation (Borgman, 2015).

This article explored the complex intersection of algorithmic bias and intellectual freedom within AI-driven digital libraries. Without deliberate ethical interventions, there is a real danger that AI technologies will reinforce social inequalities and contract the breadth of knowledge libraries endeavour to maintain. Mitigating these risks requires purposeful strategies: integrating diverse and inclusive datasets; designing AI systems that are transparent and explainable; enhancing user agency by offering greater personalisation and control; and instituting ethical oversight at the institutional level. By embedding ethical and human-centered values into technological innovation, digital libraries can embrace AI advancements while continuing to uphold their vital role as accessible and democratic spaces for knowledge dissemination. Preserving intellectual freedom and promoting equitable access in the AI age necessitate collaboration among librarians, technologists, and policymakers.

Artificial Intelligence (AI)-Driven Curation in Digital Libraries

Artificial Intelligence curation encompasses the application of artificial intelligence to automate the classification, recommendation, and organization of library content. Utilizing methods such as supervised learning, clustering, deep learning, and NLP, AI systems analyze user interactions, interpret metadata, and tailor resource suggestions to individual user needs (Murphy et al., 2021). These technologies seek to optimize the user experience by delivering highly personalized information retrieval. Recommendation engines, now a staple in many digital libraries, draw upon users' browsing histories, search behaviors, and inferred preferences to propose books, articles, and multimedia resources (Beel et al., 2016). Although personalization can significantly enhance user satisfaction and resource discoverability, it also raises pressing ethical concerns regarding the restriction of intellectual exploration and the perpetuation of existing biases (Noble, 2018).

A further ethical challenge is the opacity associated with many AI models, especially those driven by deep learning algorithms. The "black box" characteristic of these systems means that neither librarians nor users can easily interpret or question the logic behind specific curation decisions (Doshi-Velez & Kim, 2017). This lack of transparency endangers foundational library principles such as neutrality, accountability, and equitable access. Therefore, critical evaluation and oversight of AI-powered curation systems are imperative for academic and public libraries striving to maintain their commitment to open, unbiased information services.

Algorithmic Bias in AI-Driven Curation

Defining Algorithmic Bias

Algorithmic bias described consistent and systematic errors within artificial intelligence (AI) systems that result in unjust outcomes, often disadvantaging specific individuals, groups, or perspectives. Such biases typically arose from imbalanced or incomplete training data, flawed feature selection, or the uncritical reproduction of existing societal and historical inequalities embedded in source materials (Barocas, Hardt, & Narayanan, 2019; Mehrabi et al., 2021). Within the context of digital libraries, these biases can profoundly influence both the accessibility and diversity of information. Marginalized authors and subjects, for instance, may be persistently overlooked, thereby constraining the intellectual breadth available to users (Noble, 2018). At the same time, prevailing cultural narratives may be unduly amplified, fostering a misleading sense of uniformity or consensus. Additionally, AI-based recommendation systems can unintentionally perpetuate harmful stereotypes tied to gender, race, or socio-economic background, reinforcing rather than dismantling systemic inequalities (Binns, 2018). Addressing algorithmic bias is therefore essential to safeguarding intellectual freedom and promoting equitable access to diverse knowledge within AI-enhanced library platforms.

Sources of Bias

Data Bias

Biases rooted in historical and cultural contexts within training datasets play a major role in producing discriminatory behaviours in AI systems, thereby deepening existing societal inequalities (Mehrabi et al., 2021). When the content used to train these models primarily reflects dominant cultural narratives such as those authored by Western, male figures, AI technologies inevitably replicate and amplify those biases. For instance, digital library collections heavily weighted toward works by Western male authors tend to cause recommendation algorithms to favor similar material, which systematically sidelines non-Western, female, and other marginalized perspectives (Bender, Gebru, McMillan-Major, & Shmitchell, 2021).

These problems are exacerbated by metadata conventions that often overlook the cultural depth and distinct identities of underrepresented groups (Borgman, 2015). As a result, indigenous narratives, diverse epistemologies, and alternative knowledge frameworks frequently remain invisible within digital repositories. Such erasure contributes to a monolithic understanding of knowledge, posing significant threats to intellectual freedom and democratic participation (Barocas et al., 2019). Tackling these concerns necessitates proactive strategies, including the creation of inclusive datasets, the refinement of metadata standards to better represent cultural diversity, and the implementation of fairness-driven algorithmic practices, ensuring that digital libraries promote a more inclusive and pluralistic information landscape.

Design Bias

The design decisions made during the development of algorithmic systems play a pivotal role in both the creation and reinforcement of bias. One significant contributor to this problem is the selection of features that algorithms prioritised when recommending information. For instance, when recommendation engines assign considerable importance to citation counts as measures of quality, they inherently favor established academic works and prominent scholars, often at the expense of newer or marginalized voices (Mehrabi et al., 2021). Citation behaviours themselves frequently mirror the biases of historically dominant academic groups, leading to the persistent underrepresentation of minority researchers, interdisciplinary studies, and non-Western viewpoints (Barocas et al., 2019). Similarly, an excessive reliance on numerical indicators such as download counts, journal impact factors, or h-index scores can further entrench these disparities, sidelining valuable but less widely recognized contributions (Beel et al., 2016).

Such biases not only influence which knowledge becomes visible but also affect the academic development of early-career researchers by shaping the works they are most likely to read and cite. Thus, it is crucial to adopt a more reflective and equitable approach to feature selection—one that explicitly incorporates diversity, inclusion, and fairness—when implementing AI systems within digital libraries (Mitchell et al., 2019).

Interaction Bias

Bias in AI-driven recommendation systems can evolve dynamically through ongoing user interactions. Algorithms that tailor content suggestions based on a user's previous behaviours tend to shape future preferences, producing a "feedback loop" that progressively narrows the

diversity of information users encounter, ultimately resulting in the formation of "filter bubbles" (Pariser, 2011). Within these bubbles, individuals are predominantly exposed to content that mirrors their preexisting interests or beliefs, while differing perspectives are systematically excluded. This process diminishes intellectual variety and curtails opportunities for critical analysis and broader inquiry (Nguyen et al., 2014).

Moreover, personalisation features in digital libraries often elevate widely popular or mainstream resources at the expense of unconventional or minority viewpoints (Bozdog, 2013). As users increasingly interact with algorithmically recommended materials, these systems perceive such engagement as validation of their accuracy, thereby deepening patterns of biased curation (Sunstein, 2018). Over time, AI platforms may transition from neutral tools to influential actors that actively shape and limit the informational landscape. Confronting the biases that arise through user interaction is therefore crucial to maintaining intellectual freedom and ensuring that digital libraries remain spaces for diverse, critical, and inclusive knowledge exploration.

Intellectual Freedom and Its Importance

The Principle of Intellectual Freedom

Intellectual freedom refers to an individual's right to access, explore, and receive information from various perspectives without interference or censorship. This principle is fundamental to democratic societies, as it fosters informed citizenship, promotes critical thought, and supports personal autonomy (American Library Association, 2017). The unrestricted flow of ideas is essential for the advancement of education, scientific research, and active participation in civic life (Dudley, 2018). Libraries have long championed this right by curating collections that represent diverse viewpoints and resisting efforts to censor or exclude marginalized perspectives (Knox, 2015).

However, the increasing integration of artificial intelligence (AI) into library services introduces new ethical dilemmas. While AI tools designed to optimised user engagement can improve access to information, they can also inadvertently limit diversity by favoring mainstream or popular materials (Noble, 2018). Personalized algorithms can create "filter bubbles," confining users to familiar content and stifling the exposure to new or challenging ideas (Pariser, 2011). To preserve their core mission, libraries must critically assess and responsibly manage AI technologies, ensuring that these innovations enhance, rather than undermine, the essential value of intellectual freedom in today's digital world.

Threats to Intellectual Freedom from AI Curation

Artificial Intelligence-powered curation systems, designed to improve user experiences by offering personalized content, can unintentionally compromise intellectual freedom. A key issue lies in the limitation of information diversity. Algorithms often recommend content based on previous user interactions, leading to the creation of a "filter bubble" that restricts exposure to unfamiliar viewpoints (Pariser, 2011). This situation threatens the fundamental right to access a

broad array of perspectives, which is central to democratic values (American Library Association, 2017). Furthermore, AI systems often favor widely popular or mainstream content, which typically generates higher engagement, thereby amplifying dominant cultural narratives and sidelining minority or dissenting opinions (Noble, 2018). Such trends reinforce existing societal inequalities, stifling the diversity of thought needed for critical discourse and academic progress (Barocas et al., 2019). Over time, this bias could result in a more homogenous information environment where alternative viewpoints are consistently marginalized. To safeguard their commitment to intellectual freedom, digital libraries must proactively evaluate and reduce the biases inherent in AI-driven curation practices (Burrell, 2016).

Ethical Navigation Mitigating Algorithmic Bias Inclusive and Representative Data Sets

An important approach to minimizing algorithmic bias in digital libraries involves ensuring that the data used for training reflects a broad spectrum of perspectives. Libraries must shift from a passive approach to collection development to an active effort in curating inclusive digital collections that represent diverse cultures, languages, fields of study, and lived experiences (Noble, 2018). The design of metadata should reflect the intricate nature of these resources, avoiding simplistic labels that may overlook or marginalise underrepresented viewpoints (Borgman, 2015). Inclusive metadata practices not only improve the visibility of diverse materials but also help AI systems more accurately capture and represent the full spectrum of human knowledge.

Forming partnerships with marginalized communities, independent publishers, and global organizations is also crucial. Projects such as community-based archiving, the documentation of indigenous knowledge, and collaborative efforts in digitization contribute to a more inclusive informational foundation for AI systems, thereby ensuring a wider and fairer representation (Caswell, 2016; Duarte & Belarde-Lewis, 2015). Supporting open-access initiatives further promotes the democratization of knowledge by removing barriers imposed by commercial publishing (Suber, 2012). By embedding principles of diversity, equity, and inclusion at the data level, libraries can help mitigate algorithmic biases, while continuing their commitment to intellectual freedom and equitable access for all.

Transparent and Explainable AI

Ensuring transparency in artificial intelligence (AI) decision-making, often referred to as Explainable AI (XAI), is essential for upholding accountability and fostering user trust (Doshi-Velez & Kim, 2017). In digital libraries, where AI-driven curation directly shapes the content users engage with, transparency becomes both a technical requirement and an ethical responsibility. Libraries must prioritize AI systems that provide clear, understandable explanations for their recommendations, enabling users to question and challenge the algorithm's outputs when needed. Without transparency, biases inherent in AI models may remain unnoticed,

reinforcing existing inequities and compromising the fundamental principle of intellectual freedom (Wachter et al., 2017).

Effective strategies to enhance transparency include the creation and use of tools like model cards—concise documents outlining a model’s purpose, performance metrics, and ethical considerations (Mitchell et al., 2019), and datasheets for datasets, which offer detailed metadata on the origins, structure, and limitations of the training data (Gebru et al., 2018). These resources help librarians and developers assess AI systems more effectively and empower users to make informed choices about the information they interact with. Additionally, regulatory standards such as the General Data Protection Regulation (GDPR) in Europe highlight the importance of a “right to explanation,” emphasizing the societal demand for accountable and transparent AI systems (Goodman & Flaxman, 2017).

User Autonomy and Customization

Allowing users to regulate their level of personalised content is a crucial approach to mitigating the ethical issues associated with AI-based content curation. Providing adjustable personalization options enables users to choose the extent to which algorithmic filters influence their experience, thereby safeguarding their intellectual independence (Clement, Schade, & Obar, 2018). Rather than simply receiving recommendations based on past behaviours, users could tailor algorithmic parameters, broaden suggestion algorithms, or opt out of personalised content entirely. This approach supports the principles of informed decision-making and respects individual autonomy, both of which are foundational to intellectual freedom (American Library Association, 2017).

Moreover, digital libraries could offer "exploration modes" that prioritise chance discovery over algorithmic similarity. Presenting users with materials that fall outside their typical interests such as works from diverse fields, cultures, or lesser-known authors, helps break the cycle of filter bubbles, fostering a richer intellectual experience (Sunstein, 2017). Studies show that features encouraging serendipitous encounters enhance user satisfaction and expand exposure to a variety of viewpoints (Said, Jain, & Bellogín, 2021). By integrating options for user customization and discovery, libraries can more effectively serve as inclusive, democratic portals to a broad spectrum of knowledge.

Ethical Review and Oversight

Effective institutional oversight is essential for maintaining ethical practices in AI-driven digital libraries. Creating ethics review boards or AI-focused committees can be instrumental in ensuring that AI technologies align with fundamental values such as intellectual freedom, diversity, and equity. These oversight groups are tasked with formulating clear guidelines that emphasize responsible AI usage within libraries, covering principles like transparency, accountability, fairness, and inclusivity (Jobin, Ienca, & Vayena, 2019). In addition to developing policies, these bodies play a key role in overseeing the enforcement of these

principles through regular evaluations to identify and address biases, discriminatory behaviors, or any systemic exclusions present in the AI curation processes (Floridi et al., 2018).

Conducting ongoing audits and impact evaluations helps libraries proactively address ethical risks, while also fostering a culture of accountability and continuous improvement (Raji et al., 2020). By involving a diverse group of stakeholders including librarians, ethicists, technology experts, and members of marginalized communities ethics committees ensure that the perspectives of all groups are considered in the creation and application of AI systems (Whittlestone, Nyrup, Alexandrova, & Cave, 2019). This inclusive process not only builds trust within the community but also affirms the library's commitment to being an equitable and democratic knowledge-sharing space. As such, institutional oversight is a critical element for ensuring that AI technologies are integrated in an ethically responsible and sustainable manner.

Challenges in Ethical Implementation

Despite the best intentions, several challenges complicate ethical AI deployment in digital libraries:

Technical Complexity

A primary obstacle in integrating ethical AI into digital libraries is the lack of transparency in many machine learning models, especially those that utilize deep learning techniques. These models are often referred to as "black boxes" because they can produce highly accurate outcomes without providing clarity on the decision-making process (Burrell, 2016). The intricate nature of these algorithms makes it difficult for library staff and even AI experts to fully understand or evaluate the results of AI-driven curation. This lack of clarity undermines essential ethical principles like transparency and accountability in information services (Doshi-Velez & Kim, 2017). While Explainable AI (XAI) is making strides, there continues to be a trade-off between the interpretability of systems and their ability to perform at scale, particularly when handling large and diverse data sets typical of library collections (Adadi & Berrada, 2018).

Resource Constraints

The development and maintenance of ethically responsible AI systems require considerable investments in both technical expertise and financial resources. Many libraries, especially in lower-income regions, are constrained by limited budgets and a lack of qualified personnel capable of handling advanced AI tools (Ifijeh & Yusuf, 2020). As a result, such libraries may resort to commercial AI solutions, which prioritize efficiency over ethical considerations, inadvertently amplifying issues like bias and inequity.

Balancing Personalisation and Diversity

While AI-driven personalisation in digital libraries can enhance user engagement by recommending content based on individual preferences, it can also lead to the creation of "filter bubbles." These bubbles limit exposure to alternative viewpoints and may undermine the core principle of intellectual freedom (Bozdag, 2013). Library professionals face the ongoing

challenge of finding a balance between personalizing content and preserving diversity in the information user's encounter.

Conclusion

The integration of AI into digital libraries introduces both immense opportunities and substantial ethical challenges. On the positive side, AI can greatly improve information retrieval, resource discovery, and user engagement, making library systems more efficient and responsive. Advanced algorithms can streamline cataloging processes, customize content recommendations, and provide fast, precise access to vast collections. However, the ethical risks associated with AI, particularly regarding intellectual freedom, are significant. Algorithmic biases—whether due to faulty data, lack of transparency in decision-making, or unintended design flaws—can exclude marginalized voices, reinforce societal inequalities, and limit access to a wide range of perspectives.

In light of these risks, it is crucial for librarians, technologists, and policymakers to collaborate in ensuring that AI systems are ethically designed, deployed, and monitored. Key steps in this process include prioritizing diverse and inclusive training data, ensuring transparency in algorithms, safeguarding user autonomy, and creating robust accountability structures. Managing AI in digital libraries is not merely a technical challenge but a critical ethical responsibility. The decisions made today will profoundly influence the future of equitable access to knowledge for future generations.

References

- Adadi, A., & Berrada, M. (2018). *Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)*. IEEE Access, 6, 52138–52160.
- American Library Association. (2017). *Library Bill of Rights*. Retrieved from <http://www.ala.org/advocacy/intfreedom/librarybill>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and Machine Learning*. fairmlbook.org
- Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and Machine Learning*. fairmlbook.org
- Beel, J., Gipp, B., Langer, S., & Breitingner, C. (2016). *Research-paper recommender systems: a literature survey*. International Journal on Digital Libraries, 17(4), 305–338.
- Beel, J., Gipp, B., Langer, S., & Breitingner, C. (2016). Research-paper recommender systems: A literature survey. *International Journal on Digital Libraries*, 17(4), 305–338.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149–159.

- Borgman, C. L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press.
- Bozdag, E. (2013). *Bias in algorithmic filtering and personalization*. *Ethics and Information Technology*, 15(3), 209–227.
- Burrell, J. (2016). *How the machine ‘thinks’: Understanding opacity in machine learning algorithms*. *Big Data & Society*, 3(1).
- Caswell, M. (2016). *The Archive is Not Neutral: Designing a Community-Based Archival Project*. *The American Archivist*, 79(1), 1–28.
- Clement, A., Schade, S., & Obar, J. A. (2018). *Protecting privacy in the library: Challenges and opportunities*. *Library Quarterly*, 88(4), 416–428.
- Doshi-Velez, F., & Kim, B. (2017). *Towards a Rigorous Science of Interpretable Machine Learning*. arXiv preprint arXiv:1702.08608.
- Dudley, M. (2018). *Intellectual Freedom and Social Responsibility in American Librarianship, 1967-1974*. McFarland.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). *AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*. *Minds and Machines*, 28(4), 689-707.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). *Datasheets for datasets*. arXiv preprint arXiv:1803.09010.
- Gonzalez-Perez, J., Flores-Vivar, J. M., & Lopez-Golán, M. (2020). The Role of Artificial Intelligence in Information Retrieval: Challenges and Opportunities. *Journal of Information Science*, 46(6), 723-734.
- Goodman, B., & Flaxman, S. (2017). *European Union regulations on algorithmic decision-making and a “right to explanation”*. *AI Magazine*, 38(3), 50–57.
- Ifijeh, G., & Yusuf, F. (2020). *Covid-19 pandemic and the future of Nigeria's university system: The quest for libraries' relevance*. *Journal of Academic Librarianship*, 46(6), 102226.
- Jobin, A., Ienca, M., & Vayena, E. (2019). *The global landscape of AI ethics guidelines*. *Nature Machine Intelligence*, 1(9), 389-399.
- Knox, E. J. M. (2015). *Book Banning in 21st-Century America*. Rowman & Littlefield.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). *A survey on bias and fairness in machine learning*. *ACM Computing Surveys (CSUR)*, 54(6), 1–35.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). *Model cards for model reporting*. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229.
- Murphy, J., Khosrow-Pour, M., & Khosrow-Pour, D. (2021). *Handbook of Research on Emerging Trends and Technologies in Library and Information Science*. IGI Global.
- Nguyen, T. T., Hui, P. M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). *Exploring the filter bubble: The effect of using recommender systems on content diversity*. *Proceedings of the 23rd International Conference on World Wide Web*, 677–686.

- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Olson, H. A. (2002). *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Kluwer Academic Publishers.
- Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. Penguin Press.
- Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020). *Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing*. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33–44.
- Said, A., Jain, B. J., & Bellogín, A. (2021). *Increasing Diversity and Serendipity in Recommender Systems: A Review*. ACM Computing Surveys (CSUR), 54(5), 1-36.
- Suber, P. (2012). *Open Access*. MIT Press.
- Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). *Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation*. International Data Privacy Law, 7(2), 76–99.
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). **The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions**. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*, 195–200.
<https://doi.org/10.1145/3306618.3314289>
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). *The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions*. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 195–200.