Potential bias in AI: cultural representation and the marginalization of African art

Francesca Bignotti[1]
[1] University of Verona, Italy - francesca.bignotti@studenti.univr.it

ABSTRACT (ENGLISH)

The following paper explores the relationship between Artificial Intelligence (AI) and cultural representation, emphasizing how bias in Large Language Models (LLM) perpetuates the marginalization of African art. Building on the concept of "coloniality" and the presence of colonial structures in present-day knowledge systems, this study explores how contemporary AI systems trained on Eurocentric data sets extend and reinforce colonial hierarchies and rewrite cultural histories. The paper seeks to discuss the systemic lack of diversity and the reductionist approach towards representation of African art within digital platforms with the help of postcolonial theory, Digital Humanities and AI ethics. The approach used is a critical theoretical analysis and an interdisciplinary literature review, with a preliminary section that explores the possible application of the Bias Intelligence Quotient (BiQ) framework to the evaluation of cultural biases in the representation of African art. Thus, the paper seeks to provide mitigation techniques to reduce epistemological imbalance and encourage proper cultural inclusion of AI in promoting diversity and justice in the digital world.
Keywords: Artificial Intelligence (AI); Bias; Discrimination; Ethics of the Digital Humanities; Large Language Models (LLMs)

ABSTRACT (ITALIANO)

Possibili bias nell'IA: rappresentazione culturale e marginalizzazione dell'arte africana

Il presente articolo esplora la relazione tra Intelligenza Artificiale (IA) e rappresentazione culturale, sottolineando come i bias nei Modelli Linguistici di Grande Scala (LLM) perpetuino la marginalizzazione dell'arte africana. Basandosi sul concetto di "colonialità" e sulla presenza di strutture coloniali nei sistemi di conoscenza contemporanei, lo studio analizza come i sistemi di IA odierni, addestrati su dataset eurocentrici, estendano e rafforzino le gerarchie coloniali riscrivendo le storie culturali. L'articolo intende affrontare la sistemica mancanza di diversità e l'approccio riduzionista verso la rappresentazione dell'arte africana sulle piattaforme digitali, avvalendosi della teoria postcoloniale, delle Digital Humanities e dell'etica dell'IA. L'approccio adottato è un'analisi teorica critica e una revisione interdisciplinare della letteratura, con una sezione preliminare che esplora la possibile applicazione del framework Bias Intelligence Quotient (BiQ) alla valutazione dei bias culturali nella rappresentazione dell’arte africana. Pertanto, l'articolo mira a proporre tecniche di mitigazione per ridurre lo squilibrio epistemologico e a promuovere un'adeguata inclusione culturale dell'IA per favorire diversità e giustizia nel mondo digitale.
Parole chiave: Intelligenza Artificiale (IA); Bias; Discriminazione; Etica delle Digital Humanities; Modelli Linguistici di Grandi Dimensioni (LLM).

1. INTRODUCTION

In 2001, two art curators expressed in the catalogue for one of their exhibitions that those who are not exhibited do not exist (Hassan & Oguibe 2001), highlighting visibility's crucial role in gaining acknowledgement within society. Two decades later, this concept is still relevant, although in a different context; it is not only about galleries or public places but also about artificial intelligence (hereinafter AI). In the current world where AI has permeated almost every industry, it is not enough to simply be visible to be recognized as outstanding; it also means being embedded in the algorithms that drive our present society and future. Integrating AI into specific fields has created several opportunities (Nygren and Drimmer, 2023) but, at the same time, it has raised concerns about the existing biases and historical prejudices (Stahl & Eke, 2024). Colonialism left a long-term effect known as coloniality that still affects knowledge systems and cultural values to this day (Varshney et al., 2024).[1] The term coloniality can be

[1] It is crucial to make a distinction between data colonialism and data coloniality. Data colonialism can be defined as the use of data for colonial or imperialistic purposes whereby data from minorities is collected and used in a way that benefits colonial or imperialistic projects. Data coloniality is the structures and practices that ensure the subordination of these groups even when legal measures have been taken to address such issues, thus ensuring the continuity of control and dominance over data, including the ability to make decisions about its use.

defined as the structures of power that are a remnant of colonialism; it shows how the system of racial classification was employed to reinforce the colonial order with the emphasis on the subjugation of labour, which defines contemporary social, political and economic relations in the context of neocolonialism and systemic poverty (Arias, 2013). One example that poses challenges to a decolonial perspective is that AI favours already gathered and existing data, such as European sources, which are the most available (Nygren & Drimmer, 2023). This reliance on hierarchies poses a risk of perpetuating a perspective and goes against recent efforts to challenge Eurocentric views. By relying on distorted or incomplete representations, such models often reproduce narratives that may further alienate already underrepresented groups (Djanegara et al., 2024). Large language models (hereafter LLM) are now increasingly employed in academic research, creative writing, and professional sectors (Stöhr et al., 2024); however, they can also embed bias based on their data set and could unintentionally contribute to maintaining inequalities.[2] This theoretical analysis aims to understand how biases in LLMs can marginalize African art, perpetuating colonial epistemologies (Scheuerman et al., 2021) and Eurocentric theories. Furthermore, it incorporates Digital Humanities principles to offer practical measures to help reduce and mitigate these biases and create a more inclusive and equitable future.

2. METHODOLOGY

The paper's methodology is anchored on a theoretical approach whereby the research critically evaluates the current inter-disciplinary literature from postcolonial studies, critical theory, and recent contributions to AI and LLMs' ethics and cultural implications. Through these understandings, the study develops a conceptual framework to identify the epistemological and cultural processes of representing African artistic practices in AI data sets. In addition, a preliminary experimental section was developed to illustrate how cultural biases in LLMs can be evaluated.

3. BIAS AND THE MARGINALIZATION OF AFRICAN ART IN AI SYSTEMS

An essential strength of AI is that it can process large amounts of data and identify meaningful patterns; however, this capability highlights a potentially significant limitation. The data that fuels these models are not free from bias and often paint a skewed picture of non-Western societies (Tiribelli et al., 2024). AI and LLMs biases can be identified in different categories, affecting the system's precision and, essentially, the fairness of its outputs. These biases can be attributed to issues related to the training data, algorithm design or interactions between users and systems (Ferrara, 2023b). A clear example is the "Generative bias", which appears in systems that generate content like images, text or other output forms (Ferrara, 2023a). On the other hand, "Measurement bias" occurs when training data is collected or sampled unevenly, leading to the over or under-representation of certain social groups (Ferrara, 2023a). When AI models and LLMs are provided with a dataset which is prejudiced or incomplete, then the output generated will also be prejudiced and may amplify some cultures while reducing others (Chaka, 2022); in fact, discrimination based on culture, gender, ethnicity, occupation, or religion in AI is widely documented (Weidinger et al., 2021). When it comes to art history, for instance, LLMs have been trained on a mostly Western canon and, therefore, tend to favour European or American traditions (Mhlambi & Tiribelli, 2023) at the expense of non-Western art, predominantly African cultural heritage. This oppression is not just a question of data bias but a question of colonialism that, to this day, influences the process of categorization and valuation of culture across the globe (Baradaran, 2024). The exclusion of African art in AI models reflects the same colonialism (Tiribelli et al., 2024) that impacts many aspects of life, including cultural assets. Western art has long been privileged as the standard of artistic excellence, while African traditions have often been relegated to the margins or excluded from mainstream discourse altogether. These practices have contributed to the perception of African artistic practices not as part of a global artistic heritage but as a separate and inferior category confined to a distant and "primitive" past (Kasfir, 1992). LLMs trained within this colonial framework can replicate these exclusionary patterns (Kizhner et al., 2021), reinforcing that African art is not integral to global art history.
Indeed, despite efforts to mitigate biases in AI, responses generated by LLMs may reflect a generic tone when addressing African cultural heritage; in attempts to compensate for the historical

[2] Large Language Models are constructed through the formation of statistical-probabilistic connections between tokens, which represent the smallest textual units into which words are broken down during the model's training phase. For this reason, if an incomplete or decontextualized prompt is provided, the system will adapt it in the output using the terms that most frequently co-occur with those used by the user.

underrepresentation of African artistic traditions, these models often rely on generic statements, sometimes overly simplistic (Gururangan et al., 2018).
One reason for this knowledge gap is linguistic hegemony (Zeng & Yang, 2024), defined as one language's dominance over the other in social, cultural, and technological aspects. In technology and AI, languages such as English dominate data collection, algorithm development and content creation (Ferrara, 2023b). This results in a disparity whereby speakers of languages that are not dominant are underrepresented, their cultures are depicted less, and their opinions are shaped or not heard at all by the dominant linguistic group. This linguistic dominance affects the perception of people of different languages as well as the perception of their cultures on a global level. AI systems like LLMs are trained on data reflecting this linguistic hegemony, with datasets sourced mainly from the internet, where about two-thirds of the content is in English (El Louadi, 2024). As a result, the data from the developed countries is abundant, while the data from the developing countries, mainly from Africa, is scarce; for instance, out of the 54 African countries, datasets from Egypt were only used for machine learning models between 2015 and 2020 (El Louadi, 2024). This means that the African continent is portrayed through the representations of others; Africans are reduced to consumers of information about their culture instead of being producers of knowledge. Africa has a long way to go in adopting emerging technologies such as AI and meeting the increasing need for digital skills and digitization (Whaanga et al., 2015). In sub-Saharan Africa, only 6% of the population has broadband access, limiting both participation in the digital economy and representation in AI models. This lack of access, combined with the underrepresentation of Africa in digital datasets, naturally contributes to bias in AI systems (El Louadi, 2024).

4. BIAS MITIGATION FOR EQUITABLE AFRICAN ART REPRESENTATION IN AI

In Africa, digital marginalization, data marginalization, and algorithmic exclusions are major concerns that still define the current state of oppression and marginalization in the use of digital technologies. The digital technologies, big data, and algorithms used by organizations, corporations, institutions, and governments within different data jurisdictions contribute to this. Among the most significant problems are data ableism, data disablism, and data colonialism, which are evident when data generated, organized, captured, and manipulated by individuals and communities that are not in vogue is captured by big players (Chaka, 2022). This use of data simply maintains the existing structural inequalities where the value obtained from the data of powerless communities is captured for economic gain or power. Thus, data colonialism becomes one more mechanism to sustain such inequalities and is significantly damaging in the Global South, where the digital divide compounds the problems of accessibility and inclusion. As outlined, the underrepresentation of African cultural heritage in AI datasets presents a quantitative deficit, which is the lack of proportional data and a qualitative deficiency, where existing data fail to capture these artistic traditions' cultural richness fully. This dual bias results from systemic epistemological asymmetries that privilege the Eurocentric discourse, thus producing AI-generated distorted or oversimplified and generic outputs. This is because AI systems rely on probabilistic approximations, and without all the information, it tends to make content disconnected from specific knowledge. In these cases, models generate what are commonly referred to as hallucinated outputs (Banerjee et al., 2024), artificially fabricated and vague information presented as truth, often with a superficial and generic tone.[3] While these hallucinated results are difficult to eliminate entirely (Banerjee et al., 2024), some solutions can be implemented to mitigate such biased responses.
Addressing bias requires interventions targeting both data quantity and quality, starting with FAIR principles (Findable, Accessible, Interoperable, and Reusable) that provide a framework for organizing, structuring and curating datasets (Raza et al., 2024). These principles include practices such as detailed documentation of data provenance, clear descriptions of the processes used for collection and annotation and the mechanisms for updating the data regularly to capture changes in culture and technology (Huerta et al., 2023). The problem of bias in AI systems is not only a technical one but a social one, which requires a change in attitudes towards diversity (Schwartz et al., 2020) and international (Foka & Griffin, 2024) as well as interdisciplinary collaboration (Morris, 2020). In this manner, a collaborative and community-sensitive data validation process through collaborative annotations, contextualization, and enhancements ensures that cultural knowledge is dynamic and accurate. By integrating specific expertise into every phase of the AI workflow, systems can generate more nuanced outputs and avoid perpetuating hollow discourses that present artificial parity.

[3] Cambridge Dictionary, s.v. "hallucinate," (Accessed 27/12/2024), https://dictionary.cambridge.org/dictionary/english/hallucinate.

Including non-Eurocentric archives and African museum collections in AI training datasets (Costanza-Chock, 2020) is an essential measure for promoting diversity; therefore, traditional African artefacts, texts, and other digital resources from African institutions should be incorporated into the datasets. Digitizing artefacts and recording data histories preserved in African institutions ensures that datasets reflect diverse cultural perspectives, challenging the structural dominance of Western epistemologies. While digital content from Africa exists (Duarte, 2021), the continent's contribution to global data creation remains minimal. Africa's role in the global digital landscape is relatively marginal, with limited involvement in web domain registration and AI research (Ojanperä et al., 2017). Despite a broad internet user base, Africa lags in data infrastructure, with significantly fewer data centres than regions like the United States and Europe. Much of the data generated in Africa is often stored outside the continent, restricting its ability to leverage its digital potential fully (El Louadi, 2024). However, merely increasing data volume is insufficient; creating culturally specific data in collaboration with art historians (Nygren & Drimmer, 2023), cultural heritage experts and Digital Humanities scholars is equally crucial. This process ensures that AI systems generate outputs preserving African art's depth, diversity, and interpretative integrity rather than perpetuating hollow approximations (Foka & Griffin, 2024).
In addition to developing diverse data sets, it is also important to establish measures for real-time bias detection and mitigation in AI systems (Raji et al., 2020). Bias detection tools and fairness auditing protocols allow continuous monitoring of datasets and model outputs. These tools provide quantitative and qualitative scores to identify underrepresentation, stereotypes or hallucinated content, triggering immediate corrective actions. For example, outputs can be compared and cross-checked against annotations reviewed and maintained by field experts (Foka & Griffin, 2024) to identify recurring biases or Eurocentric tropes. To ensure transparency (Mitchell et al., 2019), there must be records of all the decisions that have gone into collecting data, annotating it and training the model.
A general critical decolonial approach has already been suggested to deal with datafication, algorithms, and digital citizenship from a combination of critical approaches integrated with the decolonial framework. This approach, named the Critical Southern Decolonial (CSD) framework, incorporates various theoretical lenses to paint an interconnected picture (Chaka, 2022). The framework supports the creation of new paradigms that can preserve the variety and plurality of the human and cultural experience and challenges the supremacy of the Western and American data and algorithm systems. Furthermore, the CSD framework also looks at the colonial dimensions of digital citizenship, where the latter is defined and practised in accordance with Eurocentric archetypes and thus does not acknowledge forms of citizenship, including Indigenous, nomadic, migrant or transnational citizenship. Instead, it proposes a new approach to digital citizenship that acknowledges diverse communities and their heterogeneity regardless of geographical location.
The increasing consciousness of epistemological inequalities and cultural erasure in datasets has led to new initiatives to fight for equality and properly represent non-Western societies. Projects such as Artificial Intelligence for Development in Africa (Siminyu et al., 2021) exemplify how digital technology can be a tool for cultural empowerment and epistemic justice. The AI4D Africa project, aimed at facilitating technological growth in the African continent, involves the creation of datasets that encompass the local languages, cultures, and contexts often excluded from global AI platforms. This project underscores how digital practices can evolve into tools for cultural equity, shifting the focus from data quantity to contextual quality and agency. Another significant project uniting the artistic context, AI, and its related biases and gaps was the An[0]ther {AI} in Art Summit (New Museum, 2019)[4]; which, for the first time, brought together 80 artists, scholars, curators, and technologists to address challenges and opportunities at the intersection of art and AI. Founded by Amir Baradaran, the summit was co-hosted by the Knight Foundation, Columbia University School of Engineering, New Museum, and NEW INC, exploring how AI is changing the art world and promoting equity (Baradaran, 2024). Thus, such projects prove the importance of a cross-disciplinary approach and ethical cooperation (Foka & Griffin, 2024) in utilizing digital tools to address bias. In a context where AI models often reinforce distorted representations, similar projects offer a way to imagine a future in which diversity is included and prized.

[4] Goldsmiths, University of London. (2019, May 10). An0ther AI in art: Decolonizing artificial intelligence (AI) and the future of art-making. ICCE Communication. Retrieved from https://sites.gold.ac.uk/iccecommunication/2019/05/10/an0ther-ai-in-art-decolonizing-artificial-intelligence-ai-and-the-future-of-art-making/ (Accessed 27/12/2024).

5. THEORETICAL ADAPTATION OF BIQ

Considering potential distortions and biases in the representation of African art within LLMs, this work proposes an adaptation of the Bias Intelligence Quotient (BiQ) (Narayan et al., 2024) within the context of the Comprehensive Bias Neutralization Framework (CBNF), to more precisely analyze the specific cultural biases inherent to this field of inquiry. The BiQ was originally developed as a metric capable of analyzing multiple aspects of bias, including not only the explicit content of responses but also the diversity of training data, linguistic tone, contextual sensitivity, and the model's ability to mitigate and update its outputs. This structure makes the framework particularly suitable for contexts such as African art and non-Western heritage, where epistemic flattening and Eurocentric interpretations risk being reproduced even by formally neutral models.

However, applying the BiQ in this context requires a targeted adjustment of its components. Specifically, it is proposed that the bias score be assigned a specific weight of 0.3, reflecting the centrality of explicit forms of generalization or simplification in the representation of African art. The penalty for low dataset diversity (P(d)) is understood as a fixed indicator linked to the quality and origin of the training data: in the original study, this is proposed as 0.3 for Latimer AI, which incorporates historical and cultural sources focused also on minority cultures, and 0.2 for ChatGPT, whose dataset is broader but less targeted in training.[5] Sentiment bias (s), calculated through automated semantic analysis using TextBlob, aims to detect polarization and is modulated through a coefficient λ equal to 0.5, due to its significant weight in the implicit perception of content. Context sensitivity is considered with a baseline value of 0.5 but is increased to 0.6 for prompts that directly involve the themes of colonialism, decolonization, or the restitution of African heritage.

In these cases, the model's ability to correctly interpret the historical and political framework is essential to avoid distortions on sensitive topics. In line with the original framework, it is also deemed important to evaluate the effectiveness of mitigation strategies (M) and adaptability (A). The former assesses the extent to which a model improves its responses when presented with more precise prompts or contextual terminology; the latter reflects flexibility in absorbing new trends, debates, and critical references, such as the growing role of the diaspora or practices of restitution concerning non-Western artifacts and art.

Following these assumptions and selected weights, a preliminary empirical exploration was later conducted to test the applicability of the framework to a limited number of sample queries related not only to African art tout court but also to its relationship with Western art and the themes of colonialism. The observed results, though not generalizable due to their exploratory nature, aligned with those reported in the original study, reinforcing the validity of the methodological proposal and suggesting that BiQ, if properly calibrated, can offer a solid basis for the critical analysis of bias in LLMs.

Nonetheless, it is important to emphasize the need for systematic and extended replications of this study, as discussed in Appendix 2 of the original report (Narayan et al., 2024), which outlines the potential to extend the BiQ methodology to a wide variety of prompts, allowing for broader and more robust analysis. Ongoing comparative analyses between different LLMs, even when conducted on thematically focused topics, could clearly demonstrate that the composition and origin of sources within training datasets play a decisive role in shaping the level of implicit and explicit bias. In particular, the inclusion of non-Western sources would emerge as a key strategy for designing equitable models, implying a reconfiguration of the very concept of a dataset, where cultural variety is no longer a supplementary element but a structural condition essential to ensuring ethical and epistemological quality.

6. CONCLUSIONS

This paper has highlighted how bias in AI models is not only the result of technical choices but is also a reflection of historical and cultural inequalities in the digital information universe. The predominance of a Eurocentric perspective in data sets limits our understanding of artistic diversity and reinforces narratives that may be perceived as discriminatory. To this end, these biases need to be addressed with a deliberate effort to ensure data diversity, the involvement of specialists from different fields, bias audits, and the adoption of decolonial approaches. Furthermore, a preliminary discussion considers how the BiQ framework could be adapted to assess cultural biases in the representation of African art, providing a conceptual foundation for future empirical research. The paper also suggests that African cultural heritage

[5] Latimer AI was explicitly designed to represent a wide range of cultural and social perspectives, including those that have been historically marginalized. Its training includes not only essays and publications, but also oral traditions and localized archives. Latimer AI is named after Howard Lewis Latimer, a scientist and inventor of African descent who lived at the turn of the 19th and 20th centuries.

and other marginalized artistic histories should be included in the digital ecosystem not as a secondary addition but as an intrinsic part of the art history discipline. In pursuit of a positive change in the digital environment, it is necessary to work on mitigating biases in datasets and algorithms and appreciating the diversity of the world's artistic traditions. Thus, as was noted at the beginning of this paper, visibility is crucial for existence, ensuring representation and cultural justice is an ethical and social imperative in a world increasingly mediated by AI.

BIBLIOGRAPHY

Arias, S. (2013). Coloniality and its preoccupations. Latin American Research Review, 48(3), 214–220. http://www.jstor.org/stable/43670103

Banerjee, S., Agarwal, A., & Singla, S. (2024). LLMs will always hallucinate, and we need to live with this. arXiv. https://doi.org/10.48550/arXiv.2409.05746

Baradaran, A. (2024). Towards a decolonial I in AI: Mapping the pervasive effects of artificial intelligence on the art ecosystem. AI & Society, 39(1), 7–19. https://doi.org/10.1007/s00146-023-01771-5

Chaka, C. (2022). Digital marginalization, data marginalization, and algorithmic exclusions: A critical Southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths. Journal of e-Learning and Knowledge Society, 18(3), 83-95. https://doi.org/10.20368/1971-8829/1135678

Costanza-Chock, S. (2020). Design justice: Community-led practices to build the worlds we need. The MIT Press. https://doi.org/10.7551/mitpress/12255.001.0001

Djanegara, N. D. T., Zhang, D., Uz Zaman, H. B., Meinhardt, C., Watkins, G., Nwankwo, E., Wald, R., Kosoglu, R., Koyejo, S., & Elam, M. (2024). Exploring the impact of AI on Black Americans: Considerations for the Congressional Black Caucus's policy initiatives. Stanford Institute for Human-Centered AI. https://hai.stanford.edu/white-paper-exploring-impact-ai-black-americans-considerations-congressional-black-caucuss-policy

Duarte, C. (2021). Africa is going digital, Point of View. International Monetary Fund, Finance & Development, March 18–20. https://www.imf.org/en/Publications/fandd/issues/2021/03/africas-digital-future-after-COVID19-duarte

El Louadi, M. (2024). On the preservation of Africa's cultural heritage in the age of artificial intelligence. arXiv. https://doi.org/10.48550/arXiv.2403.06865

Ferrara, E. (2023a). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. SSRN. http://dx.doi.org/10.2139/ssrn.4615421

Ferrara, E. (2023b). Should ChatGPT be biased? Challenges and risks of bias in large language models. arXiv. https://doi.org/10.48550/arXiv.2304.03738

Foka, A., & Griffin, G. (2024). AI, cultural heritage, and bias: Some key queries that arise from the use of GenAI. Heritage, 7, 6125–6136. https://doi.org/10.3390/heritage7110287

Gururangan, S., Swayamdipta, S., Levy, O., Schwartz, R., Bowman, S. R., & Smith, N. A. (2018). Annotation artifacts in natural language inference data. arXiv. https://doi.org/10.48550/arXiv.1803.02324

Hassan, S. M., & Oguibe, O. (2001). Authentic / Ex-centric: Conceptualism in contemporary African art. Forum for African Arts, 4. https://www.proquest.com/scholarly-journals/authentic-ex-centric-at-venice-biennale-african/docview/220959008/se-2?accountid=14836

Huerta, E. A., Blaiszik, B., Brinson, L. C., Bouchard, K. E., Diaz, D., Doglioni, C., Duarte, J. M., Emani, M., Foster, I., Fox, G., Harris, P., Heinrich, L., Jha, S., Katz, D. S., Kindratenko, V., Kirkpatrick, C. R., Lassila-Perini, K., Madduri, R. K., Neubauer, M. S., Psomopoulos, F. E., Roy, A., Rübel, O., Zhao, Z., & Zhu, R. (2023). FAIR for AI: An interdisciplinary and international community building perspective. Scientific Data, 10(487). https://doi.org/10.1038/s41597-023-02298-6

Kasfir, S. L. (1992). African art and authenticity: A text with a shadow. African Arts, 25(2), 41–97. https://doi.org/10.2307/3337059

Kizhner, I., Terras, M., Rumyantsev, M., Khokhlova, V., Demeshkova, E., Rudov, I., & Afanasieva, J. (2021). Digital cultural colonialism: Measuring bias in aggregated digitized digitized content held in Google Arts and Culture. Digital Scholarship in the Humanities, 36(3), 607–640. Article fqaa055. https://doi.org/10.1093/llc/fqaa055

Mhlambi, S., & Tiribelli, S. (2023). Decolonizing AI ethics: Relational autonomy as a means to counter AI harms. Topoi, 42, 1–14. https://doi.org/10.1007/s11245-022-09874-2

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–229. https://doi.org/10.1145/3287560.328759

Morris, M. R. (2020). AI and accessibility. Communications of the ACM, 63(6), 35–37. https://doi.org/10.1145/3356727

Narayan, M., Pasmore, J., Sampaio, E., Raghavan, V., & Waters, G. (2024). Bias Neutralization Framework: Measuring Fairness in Large Language Models with Bias Intelligence Quotient (BiQ). arXiv. https://doi.org/10.48550/arXiv.2404.18276

Nygren, C., & Drimmer, S. (2023). Art history and AI: Ten axioms. International Journal for Digital Art History, 9, 5.02–5.13. https://doi.org/10.11588/dah.2023.9.90400

Ojanperä, S., Graham, M., Straumann, R., Sabbata, S. D., & Zook, M. (2017). Engagement in the knowledge economy: Regional patterns of content creation with a focus on Sub-Saharan Africa. Information Technologies & International Development, 13, 33–51. https://itidjournal.org/index.php/itid/article/view/1479.html

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & 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. https://doi.org/10.48550/arXiv.2001.00973

Raza, S., Ghuge, S., Ding, C., Dolatabadi, E., & Pandya, D. (2024). FAIR enough: How can we develop and assess a FAIR-compliant dataset for large language models' training? arXiv. https://arxiv.org/abs/2401.11033

Scheuerman, M. K., Pape, M., & Hanna, A. (2021). Auto-essentialization: Gender in automated facial analysis as extended colonial project. Big Data & Society, 8. https://doi.org/10.1177/20539517211053712

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. Communications of the ACM, 63(12), 54–63. https://doi.org/10.1145/3381831

Siminyu, K., Kalipe, G., Orlic, D., Abbott, J., Marivate, V., Freshia, S., Sibal, P., Neupane, B., Adelani, D. I., Taylor, A., Touré, A. L. I., Degila, K., Balogoun, M., Diop, T. I., David, D., Fourati, C., Haddad, H., & Naski, M. (2021). AI4D – African Language Program. arXiv. https://doi.org/10.48550/arXiv.2104.02516

Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. International Journal of Information Management, 74. https://doi.org/10.1016/j.ijinfomgt.2023.102700

Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. Computers and Education: Artificial Intelligence, 7. https://doi.org/10.1016/j.caeai.2024.100259

Tiribelli, S., Pansoni, S., Frontoni, E., & Giovanola, B. (2024). Ethics of artificial intelligence for cultural heritage: Opportunities and challenges. IEEE Transactions on Technology and Society, 1(1), 298. https://doi.org/10.1109/TTS.2024.3432407

Varshney, K. R. (2024). Decolonial AI alignment: Openness, Viśeṣa-Dharma, and including excluded knowledges. arXiv. https://doi.org/10.48550/arXiv.2309.05030

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., Isaac, W., Legassick, S., Irving, G., & Gabriel, I. (2021). Ethical and social risks of harm from language models. CoRR, abs/2112.04359. https://doi.org/10.48550/arXiv.2112.04359

Whaanga, H., et al. (2015). He Matapihi Mā Mua, Mā Muri: The ethics, processes, and procedures associated with the digitization of indigenous knowledge—the Pei Jones collection. Cataloging & Classification Quarterly, 53(5-6), 520–547. https://doi.org/10.1080/01639374.2015.1009670

Zeng, J., & Yang, J. (2024). English language hegemony: Retrospect and prospect. Humanities and Social Sciences Communications, 11(317). https://doi.org/10.1057/s41599-024-02821-z