Article

AI, Digital Humanities, and the Legacies of Colonial Power

Kostas Karpouzis *

Article

AI, Digital Humanities, and the Legacies of Colonial Power

Kostas Karpouzis

Department of Communication, Media and Culture, Panteion University of Social and Political Sciences, Athens, Greece; kkarpou@panteion.gr; `https://kkarpou.github.io/`

Abstract: This paper examines the impact of artificial intelligence (AI) on digital humanities through a colonial lens, analyzing how AI can both reinforce and challenge colonial power dynamics. AI tools in digital humanities, such as text mining and language preservation, often perpetuate Western epistemologies and marginalize non-Western perspectives due to biases in data and algorithms. Using case studies, such as the Slave Voyages database and indigenous language preservation projects, this paper highlights AI's dual role as both a potential perpetuator of colonial legacies and a tool for decolonization. It recommends inclusive AI design, community-driven data governance, and the integration of alternative epistemologies to mitigate AI's colonial biases and promote more equitable knowledge production.

Keywords: artificial intelligence; digital humanities; colonialism; bias; natural language processing

---

1. Introduction

In recent years, the intersection of Artificial Intelligence (AI) and Digital Humanities (DH) has opened new pathways for analyzing, interpreting, and preserving historical and cultural data. AI techniques, such as machine learning, natural language processing (NLP), and computer vision, have allowed researchers to tackle vast datasets—ranging from digitized texts and images to archives and metadata—with unprecedented speed and precision. These innovations are transforming the field of digital humanities by providing tools to uncover patterns, identify relationships, and make inferences that were previously impossible to discern manually. However, the integration of AI into the humanities also brings forth critical questions, particularly when viewed through the lens of colonial power dynamics and knowledge production.

From the point of view of Digital humanities, related projects often focus on processing and analyzing cultural heritage data, including historical texts, artifacts, and archives. AI techniques have proven useful for projects such as digitizing and categorizing vast colonial archives, reconstructing lost languages, or creating interactive historical maps. These applications hold great potential for advancing research and public understanding of global histories and cultures. However, as these technologies expand, it becomes crucial to scrutinize how AI-driven methodologies might reproduce colonial frameworks in their approach to data collection, analysis, and representation.

The power of AI to influence digital humanities lies not only in the scale of analysis but also in the underlying assumptions encoded in algorithms. Many AI systems are trained on datasets predominantly sourced from Western contexts, with methodologies rooted in Eurocentric epistemologies. These frameworks risk perpetuating the dominance of colonial narratives by privileging Western data and ways of knowing. Consequently, AI may unintentionally reinforce the very colonial power structures that digital humanities scholars seek to deconstruct.

This colonial perspective highlights the uneven distribution of knowledge and power that emerged from centuries of colonial rule. Scholars such as Risam [18], Chakrabarty [4], and Benjamin [2] have critiqued how colonialism shaped the production and dissemination of knowledge, marginalizing non-Western voices and perspectives in the process. This dynamic persists in the digital

realm, as digital archives, AI tools, and even the infrastructure behind these technologies are disproportionately controlled by the Global North. This has led to the rise of terms like digital colonialism and data colonialism, which emphasize the continuation of exploitative practices in the digital age.

For example, archives from colonial periods, when digitized and processed by AI, often reflect the biases of the original curators—who were often colonial administrators. AI systems, unless carefully curated, risk replicating these biases by amplifying the voices and narratives of colonial powers, while silencing indigenous and marginalized groups. This calls into question the neutrality of AI as a tool and urges a critical examination of its role in shaping how historical and cultural narratives are produced and understood [5].

In this paper, we will explore the following central questions:

- How do AI-based techniques in digital humanities reproduce or challenge colonial power structures?
- In what ways do AI-driven projects in digital humanities favor Western epistemologies and exclude marginalized or indigenous perspectives?
- How can AI be employed as a tool for decolonizing digital humanities, ensuring that non-Western voices are included in global knowledge production?

The next sections discuss the theoretical framework of AI influence and the related biases, present a number of examples of how this influence is manifested in action, and propose technique to mitigate this effect.

2. Theoretical Framework

AI is proving to be a transformative technology in digital humanities, involving a range of computational techniques designed to mimic humanor augment intelligence. The most prominent AI methods applied in this domain include:

- **Machine Learning (ML)**: A subset of AI, machine learning involves training algorithms on large datasets to identify patterns, classify information, and make predictions without explicit programming. ML in digital humanities can assist in analyzing cultural artifacts by recognizing patterns across large datasets, such as identifying trends in historical texts, categorizing images, or clustering themes across vast archives. Common techniques in ML include supervised learning (where models are trained on labeled data), unsupervised learning (pattern detection without labels), and deep learning (which uses neural networks for more complex pattern recognition).
- **Natural Language Processing (NLP)**: NLP refers to AI's ability to understand, interpret, and generate human language. This is particularly relevant in the analysis of historical documents, literature, and digitized texts. NLP is used in tasks like text mining, sentiment analysis, and automated translation. Projects involving the digitization of colonial archives often apply NLP techniques to mine vast quantities of text, identifying recurring themes or uncovering previously hidden connections within historical records.
- **Computer Vision**: AI's computer vision capabilities allow machines to interpret and categorize visual information, making this technology especially useful for analyzing and preserving visual artifacts like paintings, photographs, and sculptures [15]. In digital humanities, computer vision can be applied to recognize patterns in art history, detect visual elements in colonial-era photographs, or even reconstruct incomplete artifacts.
- **Generative AI**: Generative models, like Generative Adversarial Networks (GANs), are increasingly being used in the humanities to generate art or text based on learned data. For example, these models can be trained on a dataset of colonial-era art or texts and then generate new, derivative works that reflect historical styles or content [7].

While AI offers tremendous potential for digital humanities, it also presents challenges. Many of these techniques depend heavily on the quality and composition of the data they are trained on, which can lead to biased results if not carefully curated. In the context of colonial history, this bias may

manifest as the reinforcement of Eurocentric narratives or the exclusion of indigenous perspectives, making it essential to approach AI's use critically and ethically in this field [11].

2.1. Colonialism: Definition and Framework

Colonialism refers to the political, social, and economic domination of one group over another, typically by a state or empire. Colonialism is not only a historical phenomenon but also an intellectual and cultural practice that extends beyond the physical occupation of territories. At its core, colonialism involves the extraction of resources, labor, and knowledge from colonized lands and peoples for the benefit of the colonizer.

The coloniality of power is a concept developed by sociologist Aníbal Quijano [17] to describe the lasting effects of colonialism on global systems of power. He argues that colonialism did not end with the formal independence of colonized nations but persists in the form of global hierarchies that privilege European modes of knowledge, economic systems, and social structures. This power dynamic is often reinforced through education, law, and media, where Western knowledge systems are seen as superior and universal. Colonialism also deeply impacted knowledge production. As postcolonial scholars like Said [19] have argued, colonial powers imposed their own ways of understanding the world on the peoples they colonized, marginalizing indigenous knowledge systems and languages. This legacy persists in modern academia and in many of the digital projects that aim to preserve or analyze historical artifacts, where Western epistemologies are often privileged at the expense of non-Western ones.

In the context of this work, the rise of AI and digital technologies has introduced new forms of colonialism, often described as data colonialism and digital colonialism. These terms describe the exploitation of data and digital infrastructures by powerful entities, often based in the Global North, to exert control over information, resources, and even people in the Global South.

The concept of *Data Colonialism* was developed by Nick Couldry and Ulises Mejias [6], and refers to the ways in which data is harvested from individuals and communities—often without their knowledge or consent—and then used for profit or control by large corporations and governments. The collection of data from the Global South by entities in the Global North mirrors historical patterns of resource extraction during colonial times. For example, many AI systems require vast amounts of data to function effectively, and this data is often sourced from populations in developing countries without their explicit consent or meaningful compensation. This creates a new form of dependency, where individuals and communities in the Global South become data subjects, while the benefits of AI systems are concentrated in the hands of a few powerful actors.

A related concept, *Digital colonialism*, refers to the broader control of digital infrastructures, platforms, and technologies by companies and governments in the Global North. As with traditional colonialism, this involves the extraction of value—in this case, data and digital labor—by entities that have disproportionate control over the digital tools and platforms used globally. Scholars argue that digital colonialism perpetuates existing power imbalances, as countries and communities in the Global South are often forced to rely on digital infrastructures (such as cloud services, social media platforms, and AI tools) that are owned and controlled by foreign entities. This creates an unequal distribution of power and control over digital knowledge and resources.

These forms of colonialism are particularly relevant in the context of AI-driven projects in digital humanities. For instance, many digital archives of colonial history are housed in institutions in Europe and North America, where they are digitized and analyzed using AI tools. The narratives that emerge from these projects are often shaped by Western epistemologies and priorities, rather than by the perspectives of the people whose histories are being analyzed. This raises important ethical questions about who controls the past, who benefits from the digitization of cultural heritage, and how AI might be used to either reinforce or challenge these power dynamics.

2.2. Colonialism in Action

In this section, we explore how AI-driven projects in digital humanities reflect or challenge colonial legacies. By analyzing specific examples, we can first highlight the role AI plays in either reinforcing Eurocentric knowledge systems, before proceeding to investigate opportunities for decolonization and inclusivity.

One prominent application of AI in digital humanities involves the digitization and analysis of colonial archives. For example, large-scale projects that focus on colonial-era texts, legal documents, and administrative records have become a fertile ground for AI's text-mining capabilities. However, without critical reflection, these projects can reinforce colonial perspectives by foregrounding the narratives of colonial powers while marginalizing indigenous voices.

2.2.1. The Slave Voyages Database

One well-known example of AI-driven digital humanities is the Slave Voyages project[1]. This digital archive uses AI and data analysis techniques to map transatlantic slave trade routes, analyzing records from over 36,000 voyages that transported enslaved Africans [8]. While the project offers important insights into the scale and mechanics of the transatlantic slave trade, it also raises critical questions about representation and the limits of data. The archive relies heavily on colonial records, often written by the enslavers themselves, and while it includes some data on African perspectives, it inherently prioritizes the colonizers' viewpoint, since the records were produced by those in power.

Colonial Analysis: This case reveals how AI-driven digital humanities projects, when built on colonial archives, can unintentionally perpetuate colonial worldviews. The data used in the Slave Voyages project is derived from a system that documented people as property, reflecting the dehumanizing logic of the time. As such, while the database offers an important historical resource, it also risks erasing the agency and narratives of the enslaved unless complemented with efforts to include oral histories, indigenous archives, and other non-colonial perspectives.

2.2.2. Google's AI for Indigenous Language Preservation

AI has also been applied in efforts to preserve endangered languages, many of which are spoken by indigenous communities. For example, machine learning and NLP technologies have been used to digitize and revitalize languages that are on the brink of extinction, thereby offering a potential tool for decolonizing knowledge production. However, these projects must navigate ethical challenges, particularly around ownership, consent, and the right to control indigenous knowledge.

Google's AI initiatives to preserve endangered languages, such as the use of its platform to digitize and document indigenous languages, are framed as decolonial efforts aimed at cultural preservation. However, critical voices have pointed out that this form of digitization can lead to new forms of appropriation and control, where the ownership of indigenous knowledge is shifted to global tech corporations. Indigenous communities have expressed concern over who owns the data once digitized, and whether they retain control over how their language and cultural knowledge are used.

Colonial Analysis: While AI offers powerful tools for language preservation, projects like these can fall into the trap of digital colonialism if indigenous communities are not actively involved in the process. The act of digitizing these languages can replicate the same extractive processes that defined colonialism, where the knowledge of colonized peoples was appropriated and commodified by colonial powers [21].

2.2.3. Mapping Colonial India through Image Recognition

Computer vision has been employed in projects aimed at analyzing colonial-era artworks and photographs. While this technology allows for the categorization and preservation of visual materials, it often reinforces Eurocentric notions of art and cultural heritage by prioritizing Western classification systems and neglecting non-Western art forms.

---

1   Slave Voyages: The Trans-Atlantic Slave Trade Database, Emory University, https://www.slavevoyages.org

An AI project at Oxford University sought to use computer vision to analyze photographs from colonial India. The AI system was trained to recognize and categorize images based on a dataset composed of British colonial records and photographs. While the project provided valuable insights into British colonialism's visual culture, it also raised ethical concerns. By focusing on colonial photographs, the AI system reinforced a specific narrative of India that prioritized the colonizers' gaze and treated Indian subjects as passive objects of the British Empire's control.

Colonial Analysis: This project demonstrates how AI, when trained on colonial-era data, can perpetuate colonial worldviews. By prioritizing British colonial records, the AI system reinforced a perspective that depicted Indian subjects through the lens of British control [16], rather than offering a more nuanced or indigenous understanding of these images.

2.2.4. The Colonial Despatches Project

Digital archives are a central focus of AI applications in digital humanities, offering a wealth of material for text mining and analysis. However, many of these archives consist of colonial records that reflect the biases of their creators. Without critical intervention, AI tools that analyze these archives may unintentionally reinforce the same colonial ideologies that shaped the original records.

The Colonial Despatches project [2], led by scholars in Canada, digitized and made publicly available a collection of despatches (official communications) from British colonial administrators to the British government during the 19th century. The project applied AI techniques such as text mining to analyze the language used by colonial officials in their governance of indigenous populations in British Columbia. While the project has contributed significantly to understanding colonial governance, it relies heavily on the voices of colonial administrators, often excluding the perspectives of the indigenous peoples affected by these policies.

Colonial Analysis: This project raises important questions about whose voices are privileged in the historical record. While AI has made it possible to process and analyze these texts at scale, the focus on official colonial documents risks perpetuating a one-sided view of history [3]. Projects like these highlight the need to complement AI-driven archival research with efforts to include indigenous oral histories and perspectives.

These case studies illustrate the dual potential of AI in digital humanities: while AI can, indeed, uncover new insights into historical data, it can also reinforce colonial narratives if not applied critically. Each of the above mentioned projects reveals how AI's reliance on colonial-era data or Western epistemologies can perpetuate colonial power dynamics, as discussed in the next section. However, with thoughtful design, AI can also provide a platform for decolonizing knowledge, amplifying marginalized voices, and challenging the dominance of Eurocentric narratives in the digital sphere.

3. AI's Role in Reinforcing Colonial Structures

Artificial Intelligence is often seen as a neutral or objective technology, but its deployment discussed above shows that it can actually reinforce colonial structures through algorithmic bias, data imbalances, and the privileging of Western epistemologies. As a result, AI systems, in their design, training, and deployment, can perpetuate historical injustices by embedding colonial perspectives into their outputs.

One of the main means to this is through algorithmic bias: this arises when AI systems—trained on datasets that reflect existing social, political, and historical biases—reproduce these biases in their outputs. In the context of digital humanities, this bias often results in the privileging of Western knowledge systems at the expense of non-Western or marginalized perspectives. AI's dependence on large, pre-existing datasets, many of which are sourced from historically dominant (often Western) institutions, means that these systems are inherently biased toward the values, languages, and worldviews embedded in those datasets. An example of this is evident in NLP (Natural Language Processing)

---

[2]   Colonial Despatches: The Colonial Despatches of Vancouver Island and British Columbia 1846-1871, University of Victoria, https://www.colonialdespatches.ca

systems, which are frequently used in digital humanities for text mining and analysis, are a prime example of how algorithmic bias can manifest. These systems are often trained on English or other dominant languages, which means they are less effective at processing texts written in non-Western or indigenous languages. Even when NLP systems are applied to digitized historical texts from colonized regions, the interpretation of these texts is often influenced by the linguistic and cultural biases of the training data, which tends to reflect Western norms and values [14].

AI systems in digital humanities often reinforce what philosopher Miranda Fricker calls *epistemic injustice* [9] —the systematic exclusion of certain groups from knowledge production. In colonial contexts, this means that the perspectives of indigenous or colonized peoples are often marginalized or erased, while the narratives of colonial powers are privileged. AI systems trained on colonial records are likely to perpetuate this dynamic unless deliberately designed to include alternative epistemologies or to counterbalance historical biases.

Besides this, AI's dependence on vast quantities of data raises critical issues related to data colonialism [5] to describe how digital technologies are used to extract and commodify data from individuals and communities—particularly in the Global South—without equitable benefit-sharing. In digital humanities, data colonialism manifests in the digitization of colonial archives and cultural artifacts, where the digital representation of these materials often ends up controlled by institutions in the Global North. In this context, large AI-driven projects that digitize colonial archives—such as the Endangered Archives Programme[3] at the British Library—are often framed as efforts to preserve global cultural heritage. However, critics have pointed out that these digitization projects can reinforce the control of Western institutions over cultural artifacts from the Global South. While the digital versions of these artifacts may be made publicly available, they remain housed in servers and systems owned by institutions in the Global North, perpetuating a colonial dynamic where Western entities maintain ownership and control over the knowledge and cultural production of formerly colonized peoples. This reflects a broader pattern of digital extraction, where the data and cultural resources of the Global South are harvested by corporations and institutions based in the Global North, mirroring historical patterns of resource extraction during colonial times, where the wealth of colonized nations was transferred to Europe and other colonial powers [1].

Another example of reinforcement of Western Epistemologies is manifested when AI systems in digital humanities rely on classification schemes, taxonomies, and ontologies that reflect Western academic traditions. These frameworks, embedded in AI algorithms, privilege certain ways of organizing knowledge and marginalize alternative epistemologies. For example, the way historical records are categorized and interpreted through AI tools often mirrors Western historiographical methods, which tend to emphasize linearity, causality, and written documentation—modes of historical inquiry that may not align with the traditions of oral or non-linear histories common in many non-Western cultures [18].

An example of this is AI-based analysis of colonial archives. Many AI systems are designed to categorize and interpret historical data according to Western scholarly traditions, which can lead to the exclusion of indigenous or alternative historical narratives [13]. Indigenous histories, which often prioritize oral traditions and cyclical understandings of time, may be poorly represented in AI-driven projects because these systems are built on Western ontological frameworks that emphasize written documentation and linear time. AI systems that rely on these frameworks perpetuate a form of epistemic dominance, where Western ways of knowing are treated as universal or superior. This leads to the marginalization of non-Western epistemologies and reinforces the historical silencing of colonized peoples.

---

[3]   Endangered Archives Programme, https://eap.bl.uk/

4. Decolonizing AI in Digital Humanities

In this section, we explore how AI can be employed to address and mitigate the colonial legacies discussed earlier. Decolonizing AI in digital humanities involves reshaping AI systems to be more inclusive of marginalized perspectives, ensuring ethical data use, and incorporating non-Western epistemologies into AI frameworks. This process requires the collaboration of technologists, humanists, and marginalized communities to redefine the goals and methods of AI applications in digital humanities.

4.1. Practical Examples

One of the first steps toward this is ensuring that the voices and knowledge systems of indigenous and marginalized communities are represented in the data and algorithms used. This can be achieved by working directly with these communities to curate datasets and develop AI models that reflect their cultural values, histories, and languages. An example project is the *Algonquian Linguistic Atlas*[4], a collaborative digital humanities project that uses AI to help preserve and revitalize the Algonquian languages spoken by indigenous groups in North America. By working closely with indigenous communities, linguists, and AI developers, the project has created a platform where speakers can record and share their languages, ensuring that these languages are preserved in a way that honors their cultural significance [10]. This project highlights how AI can be used to decolonize knowledge production by prioritizing indigenous agency and control over the representation and use of their cultural data. Unlike many Western-centric projects, the Algonquian Linguistic Atlas allows the communities themselves to determine how their language and culture are represented digitally.

Another de-colonization opportunity stems from the fact that decolonizing AI requires the inclusion of alternative epistemologies—ways of knowing that do not conform to Eurocentric frameworks. This includes recognizing the importance of oral histories, communal memory, and non-linear conceptions of time, all of which are integral to many non-Western cultures. AI tools can be adapted to respect these alternative epistemologies by incorporating ethical guidelines that prioritize fairness, inclusivity, and the right to control one's data. For this, we can look to New Zealand, where Māori scholars and leaders have been working on decolonizing data governance through frameworks like Te Mana Raraunga[5], the Māori Data Sovereignty Network. This framework emphasizes the importance of indigenous peoples having control over how their data is collected, stored, and used, ensuring that it serves their community's needs and values. AI tools developed under this framework must comply with the ethical principles outlined by the Māori community, such as respect for collective ownership and the right to withdraw consent at any time.

In order to decolonize AI in digital humanities, frameworks like Te Mana Raraunga must be applied more broadly. Such frameworks allow marginalized communities to have control over their own data, ensuring that AI is used in ways that benefit them rather than reinforce existing power imbalances [12]; in addition they can guide how data from colonized regions is used in digital humanities projects, ensuring that the communities from which the data originates retain control over its use and interpretation.

4.2. Challenges and Solutions for Decolonial AI

While decolonizing AI in digital humanities is an urgent task, it is not without challenges. The technical design of AI systems, the availability of data, and the need for collaboration between technologists and humanists pose significant obstacles. However, the following solutions can guide the development of decolonial AI:

---

4 Algonquian Linguistic Atlas, https://www.atlas-ling.ca
5 Te Mana Raraunga: Māori Data Sovereignty Network, https://www.temanararaunga.maori.nz

4.2.1. Community-Led Data Curation:

Engaging marginalized communities in the process of data collection, curation, and use is critical. By ensuring that the communities from which data is sourced have a say in how that data is collected and used, AI systems can be designed to reflect their values and priorities rather than reinforcing colonial power dynamics.

4.2.2. Inclusive AI Design:

AI systems should be designed with the inclusion of alternative epistemologies. This means not only integrating diverse data but also developing algorithms that can accommodate non-Western ways of organizing and interpreting knowledge. This could involve designing NLP models that are trained on indigenous languages or developing machine learning techniques that respect oral traditions as valid forms of data.

4.2.3. Critical AI Literacy for Digital Humanists:

Many digital humanists lack the technical knowledge to engage critically with AI systems. Building capacity in AI literacy among humanists can help bridge the gap between technological and ethical considerations. This would enable scholars to recognize when AI systems are reinforcing colonial perspectives and advocate for more inclusive and decolonial approaches to technology.

5. Conlusions

In this paper, we have explored the impact of AI-based techniques on digital humanities through a colonial perspective, highlighting how AI can both reinforce and challenge colonial power structures. AI's role in digital humanities has expanded the possibilities for analyzing large datasets, preserving endangered languages, and unlocking new interpretations of historical and cultural data. However, as we have seen, these advancements are not without significant ethical and epistemological concerns, particularly regarding the reproduction of colonial biases in AI systems.

One of the central arguments of this paper is that AI systems, particularly in digital humanities, are deeply influenced by the data on which they are trained and the algorithms they use. When these datasets and algorithms reflect colonial archives and Western knowledge systems, they risk reproducing the same power dynamics and epistemic exclusions that characterized colonialism. AI projects that rely heavily on colonial-era records, for instance, may prioritize the perspectives of colonizers while marginalizing or erasing the voices of the colonized.

At the same time, we have shown that AI offers powerful tools for decolonization, especially when designed with the explicit goal of including marginalized perspectives and alternative epistemologies. Projects like the Algonquian Linguistic Atlas and Māori Data Sovereignty initiatives demonstrate that AI can be used to preserve indigenous languages and knowledge systems, empowering communities to control their data and how it is represented.

To mitigate the risks of reinforcing colonial power structures, scholars, developers, and institutions working at the intersection of AI and digital humanities can adopt the following strategies:

- **Ethical AI Development:** Develop AI systems that are critically aware of the biases in their training data and seek to counterbalance these biases by incorporating diverse datasets and voices. This can include training NLP systems on indigenous languages or ensuring that visual recognition algorithms are not skewed toward Western art and cultural norms.
- **Community-Centric Approaches:** AI projects in digital humanities should engage directly with the communities whose data is being used, ensuring that these communities have control over how their knowledge is represented and used. This includes respecting indigenous data sovereignty, following ethical frameworks like Te Mana Raraunga, and fostering collaborations that benefit all parties.
- **Interdisciplinary Collaboration:** Bridging the gap between technologists and humanists is crucial for creating more equitable AI systems. Scholars in digital humanities should work alongside AI

developers to ensure that ethical and epistemological concerns are embedded into the design of AI tools from the beginning. This interdisciplinary approach can help identify potential biases and blind spots in AI applications before they become entrenched.

- **Inclusive Epistemologies:** AI-driven digital humanities projects must move beyond Western frameworks of knowledge and embrace non-Western epistemologies. This involves recognizing the validity of oral histories, communal memory, and alternative understandings of history and culture. By integrating these perspectives into AI systems, digital humanities can help decolonize knowledge production and representation.

From a research point of view, further investigation is needed to develop robust methods for identifying and mitigating biases in AI systems, especially when applied to historical and cultural datasets, and make those mitigation strategies readily available as decolonial frameworks for AI. This work can be combined with an exploration of how alternative epistemologies can be integrated into AI systems and what ethical guidelines should govern these efforts. Finally, as AI technologies continue to evolve, long-term studies on their impact on marginalized communities, especially in relation to digital cultural heritage, are needed to assess whether AI is truly aiding decolonization efforts or simply repackaging old forms of control.

References

1. Miriyam Aouragh and Paula Chakravartty. Infrastructures of empire: Towards a critical geopolitics of media and information studies. *Media, culture & society*, 38(4):559–575, 2016.
2. Ruha Benjamin. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, 2019.
3. Nicholas Blomley. *Unsettling the city: Urban land and the politics of property*. Routledge, 2004.
4. Dipesh Chakrabarty. *Provincializing Europe: Postcolonial Thought and Historical Difference*. Princeton University Press, 2000.
5. Nick Couldry and Ulises A. Mejias. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press, 2019.
6. Nick Couldry and Ulises A. Mejias. Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4):336–349, 2019.
7. Roddy Cowie, Catherine Pelachaud, and Paolo Petta. *Emotion-oriented systems: the Humaine handbook*. Springer, 2011.
8. David Eltis and David Richardson. *Atlas of the Transatlantic Slave Trade*. Yale University Press, 2010.
9. Miranda Fricker. *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press, 2007.
10. Marie-Odile Junker and Terry Stewart. A linguistic atlas for endangered languages: www. atlas-ling. ca. In *Edulearn11 Proceedings*, pages 3366–3376. IATED, 2011.
11. Kostas Karpouzis. Plato's shadows in the digital cave: Controlling cultural bias in generative ai. *Electronics*, 13(8):1457, 2024.
12. Tahu Kukutai and John Taylor. *Indigenous data sovereignty: Toward an agenda*. ANU press, 2016.
13. Walter Mignolo. *The darker side of western modernity: Global futures, decolonial options*. Duke University Press, 2011.
14. Safiya Umoja Noble. Algorithms of oppression: How search engines reinforce racism. In *Algorithms of oppression*. New York university press, 2018.
15. Maja Pantic, George Caridakis, Elisabeth André, Jonghwa Kim, Kostas Karpouzis, and Stefanos Kollias. Multimodal emotion recognition from low-level cues. *Emotion-oriented systems: The humaine handbook*, pages 115–132, 2011.
16. Christopher Pinney. *Camera Indica: the social life of Indian photographs*. University of Chicago Press, 1997.
17. Aníbal Quijano. Coloniality of power, eurocentrism, and latin america. *Nepantla: Views from South*, 1(3):533–580, 2000.
18. Roopika Risam. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Northwestern University Press, 2018.
19. Edward W. Said. *Orientalism*. Pantheon Books, 1978.

20. Yannis Skarpelos, Sophia Messini, Elina Roinioti, Kostas Karpouzis, Stavros Kaperonis, and Michaela-Gavriela Marazoti. Emotions during the pandemic's first wave: The case of greek tweets. *Digital*, 4(1):126–151, 2024.
21. Maggie Walter. The voice of indigenous data: Beyond the markers of disadvantage. *Griffith Review*, (60):256–263, 2018.