Medienimpulse
ISSN 2307-3187
Jg. 63, Nr. 4, 2025
doi: 10.21243/mi-04-25-16

Data Colonialism and Critical Data Literacy: Media Education Perspectives on Epistemic Violence in Digital Capitalism

Christian Filk

This article argues that digital capitalism generates distinctive forms of epistemic violence through data colonialism – the appropriation of human experience as raw material for algorithmic processing and capital accumulation. Drawing on decolonial theory and critical media education, the study demonstrates how contemporary datafication replicates colonial patterns of extraction, exploitation, and epistemological domination. Through three case studies – content moderation labor, biometric surveillance systems, and large language models – the analysis reveals how algorithmic systems encode and perpetuate structural violence while marginalizing alternative ways of knowing. The article proposes critical data literacy as an emancipa-

tory educational practice that expands conventional media competence frameworks through explicit attention to power, coloniality, and structural inequalities. Concrete pedagogical scenarios illustrate how learners can develop capacities for recognizing epistemic violence, understanding data extraction's political economy, and participating in building non-extractive alternatives. The study examines institutional requirements for implementation across teacher education, curriculum development, and informal learning contexts, while presenting alternative models including Indigenous Data Sovereignty and Ubuntu philosophy that challenge surveillance capitalism's extractive logic.

Der Beitrag argumentiert, dass der digitale Kapitalismus spezifische Formen epistemischer Gewalt durch Datenkolonialismus hervorbringt – die Aneignung menschlicher Erfahrung als Rohmaterial für algorithmische Verarbeitung und Kapitalakkumulation. Aufbauend auf dekolonialer Theorie und kritischer Medienpädagogik zeigt die Studie, wie gegenwärtige Datafizierung koloniale Muster der Extraktion, Ausbeutung und epistemologischen Dominanz reproduziert. Drei Fallstudien – Content-Moderation-Arbeit, biometrische Überwachungssysteme und große Sprachmodelle – verdeutlichen, wie algorithmische Systeme strukturelle Gewalt kodieren und alternative Wissensformen marginalisieren. Der Artikel entwickelt kritische Datenkompetenz als emanzipatorische Bildungspraxis, die konventionelle Medienkompetenzbegriffe durch explizite Aufmerksamkeit für Macht, Kolonialität und strukturelle Ungleichheiten erweitert. Konkrete pädagogische Szenarien illustrieren, wie Lernende Fähigkeiten zur Erkennung epistemischer Gewalt, zum Verstehen der politischen Ökonomie der Datenextraktion und zur Partizipation an nicht-extraktiven Alternativen entwickeln können. Institutionelle Anforderungen für Lehrkräftebildung, Curriculum-Entwicklung und in-

formelle Bildung werden erörtert, ebenso alternative Modelle wie Indigenous Data Sovereignty und Ubuntu-Philosophie.

"Can the subaltern speak?" – Gayatri C. Spivak (1988)

1. Introduction: Digital Capitalism as a Media Education Challenge

The concentration of power in the hands of major internet corporations represents one of the most pressing challenges for contemporary media education (Niesyto 2017). As digital platforms have become infrastructural to social, economic, and political life, they have also become sites of unprecedented data extraction and algorithmic control. This transformation has profound implications for democratic participation, social justice, and the very epistemological foundations of knowledge production. Yet mainstream media education discourse has been slow to engage with the colonial dimensions of these developments, often remaining focused on technical competencies and individual user empowerment while neglecting structural power asymmetries.

This article proposes that digital capitalism generates specific forms of epistemic violence – the delegitimization and erasure of particular ways of knowing – that require an expansion of media education's conceptual framework through decolonial perspectives. I introduce the concept of data colonialism (Couldry & Mejias 2019) as a lens for understanding how contemporary data extraction replicates colonial patterns of resource appropriation, exploitation, and epistemological domination. Through this lens,

algorithmic systems appear not as neutral technical artifacts but as sites where colonial power relations are reproduced and intensified.

This article advances a twofold argument. First, methodologically: decolonial perspectives must become constitutive of media education theory, not merely additive. Data colonialism (Couldry & Mejias 2019) reveals how contemporary datafication perpetuates colonial patterns of resource extraction, labor exploitation, and epistemological domination – patterns that cannot be adequately addressed through conventional media competence frameworks (Baacke 1996) focused primarily on individual skills. Second, pedagogically: media education must cultivate critical data literacy as a core objective – a praxis encompassing both critical consciousness (conscientização) in the Freirean sense and practical capabilities for analyzing and resisting epistemic violence embedded in algorithmic systems. This requires moving beyond technical operation to address how data systems encode and perpetuate structural violence, whose interests they serve, and what alternative sociotechnical arrangements remain foreclosed by dominant design paradigms.

This intervention is urgent because algorithmic systems increasingly mediate access to information, resources, and social participation. When these systems embody colonial logics – privileging certain languages, knowledge forms, and bodies while marginalizing others – they function as mechanisms of epistemic violence that undermine the very possibility of democratic participation.

Media education that fails to address these dimensions risks becoming complicit in reproducing digital inequalities under the guise of digital skills acquisition.

The article proceeds as follows: Section 2 outlines the methodology and epistemic positioning of the research paper. Section 3 establishes the theoretical framework by examining digital capitalism's structural principles, data colonialism as an extractive regime, and epistemic violence in algorithmic systems. Section 4 presents three case studies demonstrating how epistemic violence manifests in content moderation, biometric surveillance, and large language models. Section 5 – the article's core – develops media education perspectives on critical data literacy, proposing concrete pedagogical approaches and institutional frameworks. Section 6 examines alternative models including Indigenous Data Sovereignty and Ubuntu philosophy. The conclusion reflects on media education's role in fostering democratic resistance to digital capitalism.

2. Methodology and Epistemological Positioning

2.1 Decolonial Research Praxis: Methodology and Critical Case Study Analysis

This article employs a decolonial research methodology that challenges dominant Eurocentric paradigms while acknowledging the positionality of the researcher. Following Thambinathan and Kinsella (2021), decolonizing research requires four core practices: (1) exercising critical reflexivity about one's own epistemological

position, (2) maintaining reciprocity and respect for the self-determination of marginalized communities, (3) embracing "Other(ed)" ways of knowing, and (4) embodying transformative praxis that extends beyond academic discourse into concrete pedagogical action.

The methodological approach integrates theoretical synthesis, critical case study analysis, and pedagogical scenario development. The theoretical synthesis draws on German critical media education theory (Sesink 2014; Niesyto 2017; Filk 2025b, 2025c; Leineweber 2024) and international decolonial scholarship (Couldry & Mejias 2019; Mohamed et al. 2020; Moyo 2020) to construct an integrative framework. This cross-cultural theoretical engagement deliberately challenges the colonial habit of treating Western theory as universal while marginalizing non-Western scholarship – a pattern that Moyo (2020) identifies as requiring systematic decolonization. By centering both German critical pedagogy and decolonial thought, this article models an approach to scholarship that resists epistemic hierarchies.

The three case studies – content moderation labor, biometric surveillance, and large language models – were selected through purposive critical sampling to illustrate different manifestations of epistemic violence across the AI supply chain. Purposive critical sampling designates a selection procedure that deliberately identifies cases representing different dimensions and manifestations of a phenomenon, thereby maximizing theoretical breadth and explanatory reach. The analysis employs a decolonial lens that

centers the experiences and testimonies of marginalized communities while subjecting dominant narratives to rigorous critique. Drawing on feminist epistemology (D'Ignazio & Klein 2020) and critical race studies (Browne 2010, 2015), the case analyses attend to how power relations shape knowledge production in algorithmic systems.

2.2 Epistemological Self-Reflexivity

Critical reflexivity requires acknowledging my own positionality as a researcher situated within Western academic institutions. While this article critiques epistemic violence, it simultaneously participates in academic knowledge production systems that have historically privileged Western epistemologies. This tension cannot be resolved but must be continually negotiated through engagement with alternative knowledge systems and sustained commitment to amplifying marginalized voices – a position that Udah (2024) articulates as central to decolonial research ethics.

The decolonial turn in technology studies (Arora et al. 2023) reveals how AI systems embed colonial epistemologies through their design, training data, and deployment contexts. Yet scholarship critiquing these dynamics often remains within elite academic circuits inaccessible to affected communities. This article attempts to bridge this gap by developing pedagogical frameworks that can be adapted across diverse educational contexts, though success in this endeavor remains an open question requiring ongoing evaluation and revision in dialogue with practitioners.

Following Sims' (2024) call for "scholar-activists," this research refuses the fiction of neutral observation. An explicit normative stance – that media education should resist epistemic violence and cultivate emancipatory data literacy – aligns with critical pedagogy's commitment to transformation rather than mere interpretation (Freire 1970). However, this activist orientation brings risks of imposing external agendas onto communities with distinct priorities and strategies. The pedagogical scenarios presented in Section 5 therefore emphasize participatory design and community self-determination rather than expert-driven implementation, drawing on principles articulated in Indigenous Data Sovereignty frameworks (Carroll et al. 2020) and Ubuntu philosophy (Mhlambi 2020).

2.3 Limitations and Scope

This theoretical intervention prioritizes conceptual synthesis over empirical fieldwork. While the case studies draw on documented evidence, they do not include original ethnographic research with affected communities. Future research should employ participatory methodologies that center community knowledge and lived experience. Additionally, the focus on media education in formal and non-formal settings may obscure informal learning processes through which marginalized communities already develop critical data literacy practices. The framework proposed here should be understood as one contribution to ongoing dialogue rather than a definitive model.

The geographical scope primarily engages European and North American contexts, with case studies drawing heavily on Global South experiences of exploitation. This asymmetry reflects both the material reality of data colonialism and the epistemological challenge of avoiding extractive research practices. The alternative models presented in Section 5 (Indigenous Data Sovereignty, Ubuntu philosophy) attempt to decenter Western frameworks, though their presentation here remains mediated through my interpretive lens. Authentic decolonization would require fundamentally different research relationships including community co-authorship and benefit-sharing – directions for future collaborative work.

3. Theoretical Framework: Digital Capitalism, Data Colonialism, and Epistemic Violence

3.1 Structural Principles of Digital Capitalism

Digital capitalism represents a historically specific configuration of capitalist accumulation organized around data extraction, platformization, and algorithmic control (Staab 2019). Unlike previous capitalist formations centered on industrial production or financial speculation, digital capitalism operates through the appropriation of human experience as raw material for profit generation. Three structural principles characterize this formation:

First, platform capitalism establishes digital platforms as intermediaries that control market access and extract rent from user ac-

tivity (Srnicek 2017). Platforms like Facebook, Amazon, and Google function as infrastructural monopolies that leverage network effects to dominate multiple sectors simultaneously (Plantin et al. 2018; Klinge et al. 2023). Their power derives not from ownership of production facilities but from control over data flows and algorithmic sorting. The 2020 US House Judiciary Committee report documented how these companies have achieved "monopoly power" by functioning as "gatekeepers of commerce and communications in the digital age," controlling essential digital infrastructure upon which countless businesses and users depend (U.S. House of Representatives 2020).

This platformization of social life constitutes what Filk (2025c) analyzes as a fundamental transformation in the organization of power under conditions of algorithmic governance. In his comprehensive examination of ideology critique in platformized societies, Filk argues that platforms function not merely as technical infrastructures but as "socio-technical dispositifs" that structure possibilities for action, perception, and critique. Critical-reflexive media education must therefore address how platform architectures encode particular ideological formations while appearing as neutral technical solutions. The challenge lies in developing pedagogical approaches that enable learners to recognize and resist the "systemic rationality" embedded in platform logics – a rationality that treats datafication as inevitable and optimization as self-evidently desirable.

Second, surveillance capitalism transforms human experience into behavioral data that feeds predictive algorithms (Zuboff 2019). Every click, search, movement, and interaction becomes raw material for extracting surplus value through prediction products sold to advertisers, insurers, and other third parties. This represents what Zuboff calls an 'unprecedented economic logic' that claims human experience as free raw material for commercial exploitation.

Third, proprietary markets emerge where platforms establish exclusive control over market access and rule-making (Staab 2019). Unlike traditional commodity markets organized through price competition, proprietary markets are characterized by deliberate scarcity creation, algorithmic price discrimination, and platform owners' capacity to unilaterally change terms of access. This enables extraction of rents far exceeding traditional profit margins.

Niesyto (2017) argues that media education must develop a critical perspective on these structural transformations rather than serving merely as a skills training program for digital capitalism. He warns that media education risks becoming an 'education industry' focused on producing platform-compatible subjects unless it maintains its commitment to critical reflection and democratic participation. This critique resonates with concerns about how educational technology often reproduces rather than challenges existing power relations (Fuchs 2023).

The challenge for media education is to help learners understand these structural principles not as inevitable features of technolo-

gical development but as historically specific power arrangements that can be contested and transformed. This requires moving beyond individual empowerment narratives to collective analysis of systemic dynamics.

3.2 Data Colonialism as Extractive Regime

Data colonialism names the contemporary process through which human life is appropriated and transformed into data for capital accumulation (Couldry & Mejias 2019). The concept draws explicit parallels between historical colonialism's territorial appropriation and contemporary platform capitalism's appropriation of social life. Just as European colonialism claimed 'empty' lands for extraction, data colonialism treats human experience as an ownerless resource awaiting capture.

Three features characterize data colonialism as an extractive regime. *First*, it involves dispossession – the separation of people from control over their data and the social relations it represents. Users are positioned as data sources rather than as co-creators of value, while platforms claim ownership over aggregated data and the insights derived from it. This dispossession is legitimated through dense terms of service agreements that few read and fewer understand.

Second, data colonialism operates through extraction – the conversion of lived experience into discrete data points that can be aggregated, analyzed, and sold. Extraction transforms qualitative social relations into quantitative data flows, stripping away con-

text and reducing human complexity to trackable variables. This process is intensely asymmetric: value flows from periphery to core as data generated in the Global South enriches corporations headquartered in Silicon Valley, London, and Beijing (Thatcher et al. 2016; Milan & Treré 2019).

Third, data colonialism involves exploitation – the appropriation of value generated through users' unpaid digital labor. Every social media post, every search query, every GPS-tracked journey contributes to the data accumulation that platforms monetize. This exploitation is obscured by the 'free' provision of services, naturalizing a system where users work without recognition or compensation.

Geographic asymmetries intensify data colonialism's colonial character. While data extraction occurs globally, value accumulation concentrates in Global North technology hubs. The infrastructure of extraction – data centers, undersea cables, satellite networks – creates what Couldry and Mejias call a 'colonial cloud' that mirrors historical patterns of resource flow from colonized territories to imperial centers. Meanwhile, the Global South provides not only raw data but also the cheap labor that enables data processing, from content moderation to data annotation.

Understanding digital capitalism as a form of colonialism has significant implications for media education. It shifts attention from individual privacy concerns to structural power relations, from technical literacy to political economy, from user empowerment to collective resistance. It also creates possibilities for soli-

darity between struggles against historical colonialism and struggles against digital exploitation.

3.4 Epistemological Critique of the Data Colonialism Framework

While the data colonialism framework provides crucial analytical purchase for understanding contemporary extraction, it requires critical examination. Recent scholarship (Arora et al. 2023; Birhane et al. 2023) reveals how colonial power matrices operate not only through overt extraction but through more subtle epistemological formations that shape what counts as knowledge, whose testimony is credible, and which futures become imaginable.

Zembylas (2025) recent work on "ethno-epistemic violence" extends beyond Spivak's formulation by revealing how dominant knowledge systems not only silence subaltern voices but actively constitute subjects as epistemically inferior. In AI systems, this manifests through training data that encodes Western rationality as universal, algorithmic decision-making that privileges quantifiable over qualitative knowledge, and evaluation metrics that measure "accuracy" against datasets reflecting dominant group experiences. The violence is not merely in exclusion but in the forced translation of diverse ways of knowing into formats legible to computational systems designed around particular epistemological assumptions.

This theoretical framing itself, however, risks reproducing hierarchies it critiques. Articulating data colonialism through Western academic discourse potentially recenters European philosophical

traditions (even critical ones) while marginalizing non-Western frameworks. The concept draws on postcolonial theory (Spivak, Quijano) and critical political economy (Marx, Polanyi) – intellectual traditions emerging from specific historical-geographical contexts. While these provide powerful analytical tools, they do not exhaust possible ways of understanding datafication's violence.

Indigenous scholars propose alternative frameworks centering relationality, interdependence, and collective responsibility (Carroll et al. 2020). Ubuntu philosophy emphasizes "I am because we are" (Mhlambi 2020) – a fundamentally different ontological starting point than Western individualism undergirding both liberal privacy frameworks and Marxist alienation critiques. These alternatives do not simply 'add' perspectives but potentially reconfigure the entire theoretical apparatus for understanding technology-society relations.

Media education scholarship must navigate this tension: employing critical theory's analytical power while remaining vigilant against its potential for epistemic imperialism. Moyo (2020) emphasizes that this requires awareness of what he terms the 'politics of positionality' – a recognition that Western critical frameworks, however analytically powerful, represent perspectival rather than universal approaches to knowledge production. The following analysis of epistemic violence in algorithmic systems thus proceeds with epistemological situatedness, understanding itself as a provisional intervention in ongoing scholarly dialogue rather than definitive theoretical closure.

3.5 Epistemic Violence in Algorithmic Systems

Epistemic violence refers to the harm done through the delegitimization, erasure, or appropriation of particular ways of knowing (Spivak 1988). In the context of algorithmic systems, epistemic violence operates through the encoding of dominant epistemologies into technical infrastructure, the marginalization of alternative knowledge systems, and the violent imposition of particular ways of categorizing and knowing the world.

Spivak's original formulation analyzed how colonial discourse produced the colonized subject through epistemic frameworks that denied the validity of indigenous knowledge. The colonized could not 'speak' not because they lacked voice but because colonial epistemic structures rendered their speech inaudible. This analysis extends powerfully to algorithmic systems that embed particular ontological and epistemological assumptions while claiming universal applicability.

Quijano's (2000) concept of the coloniality of power describes how colonial hierarchies persist beyond formal decolonization through racialized classification systems that organize knowledge, culture, and labor. This coloniality operates in algorithmic systems through training data that encodes historical biases, through categories that naturalize racial and gender hierarchies, and through optimization functions that privilege Western epistemic frameworks.

Mohamed et al. (2020) introduce the concept of algorithmic coloniality to describe how AI systems reproduce colonial power rela-

tions. They develop a taxonomy of decolonial foresight encompassing three interconnected forms: algorithmic oppression (the unjust subordination of social groups through automated systems), algorithmic exploitation (the unfair extraction of value from people's data and labor), and algorithmic dispossession (the centralization of power and resources in fewer hands). Facial recognition systems that fail to detect darker skin tones exemplify algorithmic oppression, while hate speech detection algorithms that classify Black and queer vernacular as toxic demonstrate how algorithmic systems impose particular ontological frameworks as normative. Recruitment tools that discriminate against women further illustrate how algorithmic systems reproduce existing hierarchies and concentrate institutional power.

Filk's (2025a) examination of 'posthuman action orientation' provides a crucial theoretical extension for understanding how algorithmic actors redistribute agency in ways that obscure human responsibility while amplifying structural violence. Drawing on actor-network theory and posthumanist philosophy, Filk demonstrates how algorithmic systems constitute hybrid actors that exercise distributed agency across human-machine assemblages. This distribution of agency has profound implications for media education: learners must develop capabilities for analyzing how algorithms restructure the conditions of possibility for action and accountability. When predictive policing algorithms determine which neighborhoods receive enhanced surveillance, or when content recommendation systems shape information ecologies,

agency becomes diffused across socio-technical networks in ways that resist traditional frameworks of individual responsibility and democratic control.

For media education, recognizing epistemic violence in algorithmic systems requires extending Baacke's (1996) influential framework of media competence. Baacke identified four dimensions: media criticism, media knowledge, media use, and media design. His concept of media criticism already incorporates an ethical dimension that coordinates "analytical thinking and reflexive reference as socially responsible" (Baacke 1997) – *a foundation that critical data literacy can deepen by foregrounding questions of power and epistemology in algorithmic contexts:*

- Which knowledge forms are privileged or marginalized by particular algorithms?
- How do algorithmic systems construct truth and authority?
- Whose interests are served by particular epistemic frameworks?
- How can alternative epistemologies be strengthened and protected?

This expansion aligns with critical pedagogy's emphasis on conscientização (Freire 1970) – the development of critical consciousness about oppressive systems. It also resonates with Frankfurt School critiques of instrumental reason (Horkheimer & Adorno 1947), which analyzed how Enlightenment rationality dialectically generated new forms of domination. Algorithmic systems intensify this dialectic, promising rational optimization while encoding and amplifying existing power relations.

4. Manifestations of Epistemic Violence: Three Case Studies

This section examines three domains where epistemic violence manifests concretely in digital systems: content moderation, biometric surveillance, and large language models. Each case study demonstrates distinct but interconnected dimensions of data colonialism and algorithmic coloniality, while highlighting implications for media education practice.

4.1 Content Moderation: Invisible Labor and Global Exploitation

Content moderation – the process of reviewing and removing harmful content from platforms – epitomizes data colonialism's labor exploitation. In 2022, TIME Magazine journalist Billy Perrigo exposed Facebook's outsourcing of content moderation to Sama, a firm employing workers in Kenya for $1.50-2.20 per hour to review graphic violent content including murders, suicides, and child sexual abuse. The investigation documented systematic union suppression after worker Daniel Motaung attempted to organize, and Facebook's subsequent termination of its contract with Sama – leaving workers unemployed without addressing their grievances (Perrigo 2022). A 2024 medical assessment filed as part of ongoing litigation found that 144 content moderators were diagnosed with PTSD, with 81 % suffering from severe forms of the disorder (Stockwell 2024).

This arrangement exemplifies several dimensions of epistemic violence. First, it renders essential platform labor invisible to users in the Global North, who experience 'clean' feeds without awaren-

ess of the traumatic work enabling that cleanliness. Second, it positions workers in the Global South as disposable processors of content deemed too disturbing for workers in platform headquarters. Third, it extracts value from workers' psychological capacity to process violence while denying them the protections, compensation, and recognition afforded to employees at Meta's California offices.

Roberts (2019) analyzes how commercial content moderation creates precarious labor conditions for platform workers globally. Gray and Suri (2019) document in "Ghost Work" how this labor represents "the last mile" of automation – human workers performing the most complex and distressing moderation tasks that algorithms cannot yet manage. These workers must internalize platforms' content policies, which encode Western cultural norms about acceptable speech and imagery, while reviewing content from diverse global contexts (Roberts 2019; Caplan 2018). The resulting cultural violence – the imposition of one culture's norms as universal standards – compounds the direct harm of exposure to traumatic material, often leading to psychological trauma including PTSD among content moderators who review extreme violence and abuse (Roberts 2019; Gray & Suri 2019; Udupa et al. 2023)

For media education, content moderation's invisible labor poses critical pedagogical questions: How can learners develop media competence that accounts for the labor relations enabling their digital experiences? What does it mean to be a responsible plat-

form user when one's enjoyment depends on others' exploitation? A critical media education approach might involve:

- Investigating content moderation's global division of labor through research projects
- Analyzing how platforms' claims of 'community self-governance' obscure exploitative labor arrangements
- Exploring alternatives such as decentralized moderation or platform cooperatives
- Developing solidarity with content moderators through support for their organizing efforts

4.2 Biometric Surveillance: Digital Coercion and Exclusion

Biometric identification systems can exemplify epistemic violence through their imposition of particular ways of categorizing bodies and identities. The UNHCR's *Biometric Identity Management System* (BIMS) represents the world's largest humanitarian biometric database. According to a 2020 UN report, over 37 million refugees have been biometrically registered with UNHCR, increasingly making biometric registration a precondition for accessing humanitarian assistance (UNHCR 2020). While framed as improving aid distribution efficiency, BIMS transforms refugees into data subjects whose bodies become interfaces for accessing rights (Ajana 2019).

India's Aadhaar system – the world's largest biometric database – has been extensively documented for its exclusionary consequences. Following the mandatory introduction of Aadhaar-based biometric authentication in Jharkhand's Public Distribution Sys-

tem in 2016, a study by Drèze, Khalid, Khera and Somanchi (2017) found significant disruptions in food access, with beneficiaries in Ranchi district receiving only 49 % of their entitled foodgrains in July and August 2016. The *State of Aadhaar Report 2017-18* documented that 2.2 % of PDS beneficiaries in rural Rajasthan were excluded monthly due to Aadhaar-related factors – extrapolating to approximately 1.2 million people per month. Between September 2017 and November 2018, the Right to Food Campaign documented at least 17 starvation deaths in Jharkhand. Of these, at least seven victims were denied social security pensions due to Aadhaar-related issues, while additional deaths resulted from ration card cancellations or biometric authentication failures. These 'failures' disproportionately affect marginalized populations: manual laborers with worn fingerprints, elderly people with deteriorating iris patterns, and indigenous communities historically excluded from state bureaucracies. What appears as technical malfunction functions systematically to reinscribe existing marginalization (Morung Express 2018).

The epistemic violence here operates through the reduction of identity to biometric data and the treatment of bodies that don't scan as anomalous. Biometric systems embody an epistemology that privileges machine-readable bodies over lived identity. They instantiate what Browne (2010, 2015) calls 'digital epidermalization' – the reduction of personhood to skin surface and iris pattern, recapitulating colonial histories of racial classification through bodily measurement.

China's deployment of facial recognition in Xinjiang province demonstrates biometric surveillance's use for ethnic control. Recent procurement documents from multiple Chinese cities reveal that surveillance systems specifically target Uyghur faces for enhanced tracking, requiring that cameras identify whether a person is or is not a Uyghur (Grzanna 2025). Companies including Huawei, Megvii, Dahua, and Hikvision have developed "Uyghur alert" features and ethnicity recognition capabilities explicitly designed to single out members of this minority group (Human Rights Watch 2019; IPVM 2020). This represents not technical neutrality but the encoding of systematic ethnic persecution into computational infrastructure.

Media education addressing biometric surveillance must help learners understand these systems' political nature. Pedagogical approaches might include:

- Analyzing biometric systems' proliferation in learners' own contexts (school security, phone unlocking, border control)
- Examining whose bodies are made legible or illegible by biometric technologies
- Investigating alternatives to biometric identification that don't reduce identity to bodily data
- Developing critical perspectives on consent when biometric registration becomes mandatory for accessing services

4.3 Large Language Models: Linguistic Discrimination and Exploited Labor

Large language models (LLMs) like ChatGPT present as universal knowledge tools while embedding specific linguistic and cultural biases that constitute epistemic violence. Bender et al. (2021) coined the term 'stochastic parrots' to describe how LLMs reproduce training data patterns without understanding, yet are positioned as authoritative knowledge sources. Their critique highlights environmental costs, encoded biases, and the illusion of meaning as core problems with scaling language models.

Recent research quantifies LLMs' discriminatory treatment of non-standard language varieties. Fleisig et al. (2024) found ChatGPT systematically discriminated against minoritized dialects, generating 19 % more stereotypes on average compared to Standard American English, while retaining only 3 % of African American English linguistic features versus 78 % for standard varieties. This demonstrates how algorithmic systems treat linguistic diversity not as richness but as deviation requiring correction – a form of what Flores and Rosa (2015) call 'raciolinguistic ideologies' that position racialized speakers as linguistically deficient regardless of their actual language use.

Tokenization – the process of breaking text into units for processing – reveals structural bias in *large language model* (LLM) architecture. Research by Petrov et al. (2023) documents stark disparities: while English text requires approximately 1.3 tokens per word, Burmese requires 11.7 to 16.9 times more tokens for equiv-

alent content, depending on the tokenizer used. This has direct consequences: API pricing based on tokens makes LLM use approximately 12 times more expensive for languages like Burmese, while token limits exclude longer text necessary for adequate context in languages with different grammatical structures (Ahia et al., 2023; Petrov et al. 2023). The technical architecture thereby encodes linguistic hierarchy (Teklehaymanot & Nejdl 2025).

LLMs also depend on exploited labor for data annotation and content filtering. Perrigo (2023) reported that OpenAI contracted Sama to employ Kenyan workers at $1.32-2.00 per hour to label violent and sexual content for ChatGPT's safety filters. These workers processed graphic material describing rape, murder, and child abuse to train systems that would refuse to generate such content – yet another instance of platforms outsourcing traumatic labor to Global South workers.

For media education, LLMs present both pedagogical opportunities and critical challenges. Educators must help learners understand:

- How LLMs' claimed universality obscures linguistic and cultural specificity
- The labor relations enabling LLM development and deployment
- Why non-standard languages and dialects are systematically disadvantaged
- How to evaluate LLM outputs critically rather than treating them as authoritative

5. Media Education Perspectives: Critical Data Literacy as Resistance Practice

This section develops the article's central contribution: a framework for critical data literacy as an emancipatory media education practice. I first revisit and expand classical media competence concepts, then elaborate specific competencies for critical data literacy, present concrete pedagogical scenarios, and discuss institutional requirements for implementation.

5.1 Expanding the Media Competence Framework

Baacke's (1996, 1997) four-dimensional model of media competence – media criticism, media knowledge, media use, and media design – remains foundational to German media education. However, addressing data colonialism and epistemic violence requires expanding each dimension:

Media criticism must extend beyond analyzing media messages to examining colonial power structures in data systems. This involves understanding data extraction as exploitation, recognizing algorithmic discrimination, and identifying epistemic violence in technical systems. Critical data literacy asks not just 'what does this message mean?' but 'whose interests does this system serve?'.

Media knowledge must include technical understanding of data extraction mechanisms, algorithmic decision-making, and the material infrastructure of digital systems. Learners need to grasp how platforms monetize data, how training data shapes AI out-

puts, and how seemingly neutral technical choices encode power relations. This knowledge should extend to global labor conditions enabling digital services.

Media use must involve reflection on personal data practices and their collective implications. Rather than individualized privacy management, this dimension emphasizes how data sharing affects not only oneself but entire communities. It includes developing practices of refusal, obfuscation, and collective data governance as alternatives to passive acceptance of extractive systems.

Media design must encompass participation in developing non-extractive alternatives. This might involve contributing to open-source projects, participating in platform cooperatives, or designing community-controlled data systems. The goal shifts from producing content for platform distribution to building infrastructural alternatives.

Filk's (2025d) comprehensive framework for strengthening digital self-determination provides crucial guidance for integrating these expanded competencies into teacher education. His analysis demonstrates how effective media education in the age of AI requires not merely adding "AI literacy" as an additional competency domain, but fundamentally reconceptualizing the relationship between media ethics, technological understanding, and pedagogical practice. Filk argues that digital self-determination – the capacity for autonomous, informed decision-making in datafied environments – must become the normative horizon orienting all media education efforts. This requires teacher education pro-

grams that cultivate what he terms "reflexive technological competence": the ability to critically analyze technological systems while developing practical capabilities for ethical engagement with algorithmic environments. Crucially, Filk's framework emphasizes that digital self-determination cannot be realized through individual empowerment alone but requires collective capacity-building and institutional transformation.

Beyond these expansions, critical data literacy adds an epistemic dimension that explicitly addresses knowledge, power, and coloniality:

- *Recognizing*: Which knowledge forms are privileged or marginalized by particular algorithmic systems?
- *Understanding*: How do algorithms construct truth and authority? What epistemological assumptions structure data collection and processing?
- *Acting*: How can alternative epistemologies be strengthened and protected? What practices of epistemic resistance are available?

This epistemic dimension draws on decolonial theory's insistence that knowledge is always situated and that dominant knowledge systems maintain power through claiming universality. It also connects to data feminism's principle that data systems embody particular standpoints that can be challenged and transformed (D'Ignazio & Klein 2020).

5.2 Critical Data Literacy as Educational Goal

Critical data literacy synthesizes critical pedagogy (Freire 1970), data literacy education (Tygel & Kirsch 2016), and decolonial epistemology into a comprehensive framework. Following Freire, it emphasizes conscientização – developing critical consciousness about oppressive systems – while recognizing that consciousness alone is insufficient without practical capabilities for intervention. The framework distinguishes three interconnected competence domains:

5.2.1 *Knowledge* (Know-That):

- Understanding data extraction mechanisms and business models of platform capitalism
- Knowledge of global inequalities in data economics (flows, infrastructure, labor conditions)
- Awareness of epistemic violence in algorithmic categorization and classification
- Historical understanding of colonialism's relationship to contemporary data extraction

5.2.2 *Skills* (Know-How):

- Analyzing algorithmic systems for embedded biases and power relations
- Implementing data protection practices at individual and collective levels
- Identifying and using alternative, non-extractive digital tools
- Participating in collective data governance initiatives

5.2.3 Disposition (Know-Why):

- Critical reflexivity about technology and power relations
- Solidarity with communities experiencing algorithmic discrimination
- Commitment to data justice and equitable access to digital resources
- Orientation toward collective rather than individualized technological solutions

This tripartite framework – inspired by Aristotle's distinction between theoretical, practical, and ethical knowledge – recognizes that effective critical data literacy requires not only information about data systems but also practical skills for engagement and ethical commitment to justice. It moves beyond deficit models that treat learners as lacking knowledge toward recognition of their existing critical capacities and experiential knowledge.

The concept of digital sovereignty, as elaborated by Filk (2026), provides a crucial political dimension for understanding critical data literacy's emancipatory potential. Digital sovereignty refers to both individual and collective capacities for self-determined participation in digitally mediated public spheres. Filk argues that empowering youth through digital sovereignty requires moving beyond protective paternalism toward recognizing young people as political subjects capable of shaping technological futures. This involves cultivating competencies for democratic participation in debates about AI governance, data regulation, and platform accountability. Critical data literacy, from this perspective, becomes not merely a defensive practice of protection against mani-

pulation but an affirmative practice of democratic engagement. Youth must develop capabilities not only for critically analyzing existing algorithmic systems but for imagining and advocating for alternative technological configurations aligned with democratic values and social justice commitments.

5.3 Pedagogical Scenarios for Critical Data Literacy

This section presents three concrete scenarios for implementing critical data literacy education. Each targets different educational levels and emphasizes distinct aspects of the framework.

5.3.1 Scenario 1: 'Who Works for My Timeline?' (Secondary School, Ages 16–18)

Learning Objective: Students develop understanding of invisible labor in digital capitalism and reflect on their own relationship to platform labor.

Method: (1) Introduction through documentary excerpts on content moderation (The Cleaners 2018). (2) Research phase where students investigate content moderation labor conditions using TIME Magazine investigations and academic sources. (3) Interviews with peers about their platform usage and awareness of moderation. (4) Group analysis using guiding questions: What makes content moderation labor invisible? How does geographic location affect labor conditions? What alternatives exist to exploitative moderation?

Materials: Perrigo (2022) TIME article, excerpts from Roberts (2019) Behind the Screen, documentary clips, interview guides.

Critical Reflection: Discussion of personal responsibility as platform users. What does ethical social media use look like? How can users support moderator organizing efforts? Should platforms be required to employ moderators directly rather than outsourcing?

5.3.2 *Scenario 2*: 'Biometric Futures' (Secondary School/Adult Education, Ages 18+)

Learning Objective: Participants critically evaluate biometric identification systems and develop alternative visions for identity verification.

Method: (1) Mapping exercise documenting biometric systems in participants' lives (phone unlocking, border control, workplace attendance). (2) Case study analysis of Aadhaar failures and UNHCR BIMS implementation. (3) Future workshop methodology: Critique phase identifying problems with biometric systems, Fantasy phase imagining alternatives, Implementation phase developing realistic proposals for identity systems that don't reduce people to bodily data.

Materials: UNHCR BIMS documentation, Menon (2017) and Khera (2017) research on Aadhaar exclusions, Browne (2010, 2015) excerpts on digital epidermalization.

Critical Reflection: Who benefits from biometric systems? Whose bodies are made legible or illegible? What gets lost when identity is reduced to biometric data? How do biometric requirements for accessing services undermine meaningful consent?

5.3.3 *Scenario 3*: 'AI Doesn't Speak My Language' (All Levels, Adaptable)

Learning Objective: Learners recognize linguistic discrimination in AI systems and critique universality claims of language models.

Method: (1) Experimental testing with ChatGPT using different languages, dialects, or sociolects. Prompt the system identically in Standard English, African American Vernacular English, regional dialects, or non-Western languages and compare responses. (2) Documentation of differences in quality, tone, and content. (3) Analysis of Fleisig et al. (2024) and Petrov et al. (2023) study findings. (4) Creative response: students develop multilingual or dialect-specific AI prompts that resist linguistic normativity.

Materials: ChatGPT or similar LLM access, Fleisig et al. (2024) and Petrov et al. (2023) research findings, examples of linguistic discrimination.

Critical Reflection: Why do AI systems treat some languages as 'standard'? How does linguistic discrimination connect to other forms of algorithmic bias? What would multilingual AI designed from non-Western epistemologies look like? How can linguistic diversity be protected in an AI-dominated information environment?

These scenarios share several pedagogical principles: They begin with learners' lived experiences, connect individual experiences to structural analysis, emphasize collective investigation over individual assessment, and orient toward transformation rather than mere critique. They also model what Markham (2020) calls 'taking

data literacy to the streets' – embedding critical analysis in contexts meaningful to learners rather than abstract exercises.

5.4 Institutional Frameworks for Critical Data Literacy

Realizing critical data literacy requires supportive institutional conditions beyond individual educator initiative. This section outlines necessary changes across teacher education, curriculum, informal education, and research infrastructure.

Teacher Education: Pre-service and in-service teacher education must incorporate critical perspectives on digital capitalism. This requires interdisciplinary collaboration between media education, computer science, sociology, and critical theory. Teachers need not only technical skills but also frameworks for analyzing power in technological systems. Professional development should include engagement with decolonial theory, political economy of media, and pedagogies of resistance.

Filk's (2025d) model for integrating media ethics and AI into teacher training provides practical guidance for implementing these transformations. His framework proposes a three-tier approach: (1) foundational modules addressing ethical principles for algorithmic systems and their philosophical justifications, (2) subject-specific applications demonstrating how AI literacy can be integrated across curricular domains, and (3) reflective practice seminars where pre-service teachers critically examine their own technological assumptions and develop pedagogical strategies for addressing algorithmic bias. Crucially, Filk emphasizes that teacher

education must cultivate not only knowledge about AI systems but also ethical disposition and practical wisdom (phronesis) for navigating complex technological-pedagogical dilemmas in everyday school life.

Curricular Integration: Educational policy must move beyond generic 'digital competencies' to specify critical data literacy as a distinct learning objective. This requires not technical skills divorced from social context but integrated approaches that examine technology's embeddedness in power relations. Cross-curricular projects connecting media education, social studies, and ethics can model this integration.

Informal Education: Youth work and community education provide crucial spaces for experimental and participatory media pedagogy. Partnerships with hackerspaces, FabLabs, and community technology centers can enable hands-on exploration of alternative systems. These sites often have greater flexibility than schools to engage controversial topics and develop sustained relationships with learners.

Research Infrastructure: Systematic evaluation of critical data literacy interventions requires dedicated research funding and institutional support. Longitudinal studies examining outcomes beyond immediate skill acquisition – such as changes in collective organizing or platform use patterns – would demonstrate these approaches' effectiveness. Action research methodologies involving educators as co-researchers can ensure relevance to practice.

The *Vienna Manifesto for Digital Humanism* (2019) calls for digital technologies serving humanity rather than exploiting it. Realizing this vision through education requires moving beyond market-driven skills training to cultivate capacities for democratic participation and collective resistance. Media education's historical commitment to emancipatory pedagogy positions is to lead this transformation – if it maintains critical perspective rather than accommodating to digital capitalism's demands.

6. Alternative Approaches: Indigenous Data Sovereignty and Ubuntu Ethics

Beyond critique, decolonial perspectives on data offer alternative models for more just data relations. This section examines Indigenous Data Sovereignty and Ubuntu philosophy as frameworks that challenge data colonialism's extractive logic while suggesting pathways toward collective data governance.

Indigenous Data Sovereignty asserts Indigenous peoples' right to govern collection, ownership, and application of data about their communities, territories, and resources (Carroll et al. 2019). The CARE Principles – Collective Benefit, Authority to Control, Responsibility, and Ethics – provide guidelines for Indigenous data governance. These principles reject both the 'open data' assumption that all data should be freely accessible and the proprietary model where corporations own community data.

Indigenous Data Sovereignty challenges the epistemic foundations of conventional data science. It insists that data is not neutral

information but represents relationships, responsibilities, and connections to place. Data about salmon populations, for instance, is not merely environmental monitoring data but embodies relationships between peoples, species, and territories maintained across generations. Treating such data as extractable resource violates these relationships.

For media education, Indigenous Data Sovereignty offers crucial lessons: Data is never merely technical but always relational. Communities rather than individuals or corporations should govern data about collective life. Data governance requires attention to historical injustices and ongoing power relations. Alternative epistemologies – ways of knowing that don't reduce everything to quantifiable variables – must be protected rather than erased.

Ubuntu philosophy, articulated across various African contexts, emphasizes relationality, interdependence, and collective flourishing. Mhlambi (2020) applies Ubuntu principles to AI development, arguing for shifting from rationality-centered to relationality-centered approaches. Rather than optimizing individual preferences, Ubuntu-informed systems would prioritize collective well-being and mutual care.

Ubuntu's principle of 'I am because we are' directly challenges surveillance capitalism's atomized individual subject. It suggests that data about any person belongs to their community of relationships rather than to platforms or even to individuals as private property. This opens possibilities for collective data governance

where communities negotiate terms of data collection and use together rather than through millions of individual consent clicks.

Platform Cooperativism applies cooperative ownership models to digital platforms, enabling workers and users to collectively govern platform rules and share revenues (Scholz 2016). Examples include Stocksy (photographer-owned image platform), Resonate (musician-owned streaming service), and various platform co-ops for domestic workers and taxi drivers. While not explicitly decolonial, platform cooperativism demonstrates alternatives to extractive platform capitalism.

Data Cooperatives extend cooperative principles to data governance. Rather than individuals negotiating alone with corporations or surrendering data as condition for access, data cooperatives collectively bargain over data terms. The *Driver's Seat Cooperative*, for instance, enables gig workers to pool their data about platform algorithms, generating collective knowledge about how platforms manage workers while maintaining individual privacy.

These alternatives share rejection of both corporate data extraction and state surveillance. They insist on collective governance, equitable benefit distribution, and epistemic pluralism – recognition that different communities may have different values regarding data. For media education, they demonstrate that 'there is no alternative' (TINA) to current systems is false. Student-led data cooperatives, school-based platform experiments, and engagement with Indigenous data principles can model alternative practices.

7. Conclusion: Media Education Between Accommodation and Resistance

This article has argued that digital capitalism generates distinctive forms of epistemic violence through data colonialism – the appropriation of human life as raw material for algorithmic processing and capital accumulation. Three case studies demonstrated how epistemic violence operates concretely: content moderation's exploitation of Global South labor, biometric surveillance's reduction of identity to machine-readable data, and large language models' encoding of linguistic hierarchies. These are not aberrations but systematic features of how algorithmic systems reproduce colonial power relations.

Against this backdrop, media education faces a fundamental choice articulated by Niesyto (2017): accommodation to digital capitalism's demands or critical accompaniment of democratic transformation. The accommodationist path leads toward skills training that produces platform-compatible subjects – users who efficiently navigate existing systems without questioning their structure. This path may secure funding and institutional support but abandons media education's emancipatory heritage.

The alternative path – critical accompaniment – requires developing critical data literacy as core pedagogical practice. This means expanding conventional media competence frameworks to incorporate explicit analysis of power, coloniality, and structural violence. It means teaching learners to recognize epistemic violence in algorithmic systems, understand the political economy of data

extraction, and participate in building alternatives. Most fundamentally, it means treating media education not as preparation for digital capitalism but as cultivation of capacities for transforming it.

Building on Filk's (2025c) analysis of ideology critique under conditions of algorithmic governance, this article has demonstrated how critical-reflexive media education must address the distinctive challenges posed by platformized societies. The pedagogical scenarios presented here show that critical consciousness need not remain abstract but can be developed through engagement with material realities of content moderation labor, biometric systems, and linguistic AI. They also model pedagogical principles – starting from lived experience, connecting individual experience to structural analysis, emphasizing collective investigation, orienting toward transformation – that distinguish critical data literacy from technical training.

Alternative models – Indigenous Data Sovereignty, Ubuntu philosophy, platform cooperativism, data cooperatives – demonstrate that non-extractive data relations are possible. They challenge the naturalization of surveillance capitalism by instantiating different values: collective governance over individual choice, relationality over atomization, diverse epistemologies over algorithmic monoculture. Media education that engages these alternatives helps learners imagine and work toward different digital futures.

Realizing this vision requires institutional transformation. Teacher education must incorporate critical perspectives on digital capita-

lism, as detailed in Filk's (2025d) framework for integrating media ethics and AI. Curricula must specify critical data literacy as distinct from generic digital skills. Informal education spaces must support experimental and participatory approaches. Research infrastructure must evaluate not just immediate skill acquisition but longer-term impacts on collective organizing and democratic participation.

The stakes are high. Algorithmic systems increasingly mediate access to information, resources, and social participation. When these systems embody colonial logics, they undermine democratic possibility itself. Media education that addresses only technical operation while ignoring structural violence becomes complicit in digital capitalism's reproduction. But media education that cultivates critical data literacy – understanding of power, solidarity with the exploited, commitment to justice, capabilities for resistance – contributes to building more equitable digital futures.

This is not a call for media education to abandon practical skills or technological engagement. Rather, it insists that skills acquisition must be embedded in critical analysis of why particular skills are valorized, whose interests they serve, and what alternatives exist. Drawing on Filk's (2026) conception of digital sovereignty as empowerment for democratic participation, this article proposes that media education's most important contribution is not training efficient platform users but cultivating subjects capable of democratic technological self-determination – subjects who can critically analyze algorithmic power while actively participating in shaping alternative technological futures.

The path forward requires both critique and construction. We must rigorously analyze how digital capitalism generates epistemic violence while simultaneously building alternatives through collective experimentation. We must maintain unwavering attention to power while refusing cynicism that forecloses possibility. We must connect media education's local practices to global struggles against exploitation while respecting different communities' distinct needs and strategies.

Niesyto's (2017) question – accommodation or critical accompaniment – ultimately poses the question of what kind of society media education serves. If the goal is efficient integration into existing systems, technical training suffices. But if the goal is democratic transformation toward justice and equality, media education must become a practice of collective resistance. Critical data literacy offers one framework for this transformation. The work of

implementing it – in classrooms, youth centers, community organizations – now begins.

Bibliography

Ahia, Ohia/Kumar, Sachin/Gonen, Hila/Kasai, Jungo/Mortensen, David R./Smith, Noah A./Tsvetkov, Yulia (2023): Do all languages cost the same? Tokenization in the era of commercial language models, in: Bouamor, Houda/Pino, Juan/Bali, Kalika (Eds.): Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 9904–9923). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.emnlp-main.614

Ajana, Btihaj (2019): Digital biopolitics, humanitarianism and the datafication of refugees, in: Cox, Emma/Durrant, Sam/Farrier, David/Stonebridge, Lyndsey/Wooley, Agnes (Eds.): Refugee studies: Research Across the Humanities (pp. 464–488). Edinburgh: Edinburgh University Press.

Alpermann, Björn (2025, January 15): From seeing to understanding: How public surveillance targets Uyghurs with AI cameras throughout China, Remote XUAR, online at: https://www.remotexuar.com/post/from-seeing-to-understanding-how-public-surveillance-targets-uyghurs-with-ai-cameras-throughout-c (last access: 11.26.2025).

Arora, Ananya/Barrett, Michael Ian/Lee, E./Oborn, Eivor Marianne/Prince, Karl (2023): Risk and the future of AI: Algorithmic bias, data colonialism, and marginalization, in: Information and Organization 33(4), Article 100498. https://doi.org/10.1016/j.infoandorg.2023.100478

Baacke, Dieter (1996): Medienkompetenz als Netzwerk: Reichweite und Fokussierung eines Begriffs, der Konjunktur hat, in: Medien praktisch, 20(2), 4–10, actually online at: https://www.medien-paed.com/article/view/2050 (last access: 11.26.2025).

Baacke, Dieter (1997): Medienpädagogik, Hameln: Niemeyer.

Bender, Emily M./Gebru, Timnit/McMillan-Major, Angelina/Shmitchell, Shmargaret (2021): On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

Browne, Simone (2010): Digital epidermalization: Race, identity and biometrics, in: Critical Sociology 36(1), 131–150.

Browne, Simone (2015): Dark matters: On the surveillance of blackness, Durham: Duke University Press.

Caplan, Robyn (2018): Content or context moderation? Accelerating harmful speech and the dangers of computer-aided decision making, Data & Society Research Institute.

Carroll, Stephanie R./Garba, Ibrahim/Figueroa-Rodríguez, Oscar L./Holbrook, Jarita/Lovett, Raymond/Materechera, Simeon/Parsons, Mark/Raseroka, Kay/Rodriguez-Lonebear, Desi/Rowe, Robyn/Sara, Rodrigo/Walker, Jennifer D./Anderson, Jane/Hudson, Maui (2020): The CARE Principles for Indigenous Data Governance, in: Data Science Journal 18(43), 1–12. https://doi.org/10.5334/dsj-2020-043

Couldry, Nick/Mejias, Ulises A. (2019): Data colonialism: Rethinking big data's relation to the contemporary subject, in: Television & New Media 20(4), 336–349. https://doi.org/10.1177/1527476418796632

D'Ignazio, Catherine/Klein, Lauren F. (2020): Data feminism, Cambridge: MIT Press. https://doi.org/10.7551/mitpress/11805.001.0001

Drèze, Jean (2017, February 12): In Jharkhand, compulsory biometric authentication for rations sends many away empty-handed, in: Scroll.in, online at: https://scroll.in/article/829071 (last access: 11.26.2025).

Drèze, Jean/Khalid, Nazar/Khera, Reetika/Somanchi, Anmol (2017): Aadhaar and food security in Jharkhand: Pain without gain?, in: Economic and Political Weekly 52(50), 50–59.

Filk, Christian (2025a): Posthuman action orientation: Algorithmic actors and distributed agency in media education, in: Medienimpulse: Beiträge zur Medienpädagogik 63(3), 1–46, online at: https://journals.univie.ac.at/index.php/mp/article/view/9680 (last access: 11.26.2025).

Filk, Christian (2025b): Ideology critique in the age of algorithmic governance: On the transformative power of critical-reflexive media education in a platformized society, in: Medienimpulse: Beiträge zur Medienpädagogik 63(2), 1–55, online at: https://journals.univie.ac.at/index.php/mp/article/view/9469 (last access: 11.26.2025).

Filk, Christian (2025d): Strengthening digital self-determination: Integrating media ethics and Artificial Intelligence into teacher training for everyday school life, in: Medienimpulse: Beiträge zur Medienpädagogik 63(1), 1–48, online at: https://journals.univie.ac.at/index.php/mp/article/view/9263 (last access: 11.26.2025).

Filk, Christian (2026): Empowering youth through digital sovereignty: A framework for political participation in the age of AI, in: Mein, Georg et al. (Eds.): Democracy and youth in the digital age:

Evolving technologies and political participation, Heidelberg: Springer [Manuscript in preparation].

Fleisig, Eve/Smith, Genevieve/Bossi, Madeline/Rustagi, Ishita/Yin, Xavier/Klein, Dan (2024): Linguistic bias in ChatGPT: Language models reinforce dialect discrimination, in: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 13387–13401). Association for Computational Linguistics, online at: https://aclanthology.org/2024.emnlp-main.750/ (last access: 11.26.2025).

Flores, Nelson/Rosa, Jonathan (2015): Undoing appropriateness: Raciolinguistic ideologies and language diversity in education, in: Harvard Educational Review 85(2), 149–171. https://doi.org/10.17763/0017-8055.85.2.149

Freire, Paulo (1970): Pedagogy of the oppressed, London: Continuum.

Fuchs, Christian (2023): Der digitale Kapitalismus: Arbeit, Entfremdung und Ideologie im Informationszeitalter, Weinheim: Beltz Juventa.

Gray, Mary L./Suri, Siddharth (2019): Ghost work: How to stop Silicon Valley from building a new global underclass, New York: Hachette Book Group.

Horkheimer, Max/Adorno, Theodor W. (2002): Dialectic of enlightenment: Philosophical fragments (E. Jephcott, Trans.), Redwood: Stanford University Press (Original work published 1947).

Human Rights Watch (2019). China's algorithms of repression: Reverse engineering a Xinjiang police mass surveillance app, online at: https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass (last access: 11.26.2025).

IPVM (2020, December 8): Huawei tested AI software that could recognize Uyghur Muslims and alert police, online at: https://www.washingtonpost.com/technology/2020/12/08/huawei-tested-ai-software-that-could-recognize-uighur-minorities-alert-police-report-says/ (last access: 11.26.2025).

Khera, Reetika (2017): Impact of Aadhaar in welfare programmes, in: Economic & Political Weekly 52(50), 61–70.

Klinge, Tobias J./Hendrikse, Reijer/Fernandez, Rodrigo/Adriaans, Ilke (2023): Augmenting digital monopolies: A corporate financialization perspective on the rise of Big Tech, in: Economy and Society 52(1), 1–28.

Leineweber, Christian (2024): Eine kritische Bildungstheorie postdigitaler Subjektivität, in: MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung 61, 69–87. https://doi.org/10.21240/mpaed/61/2024.06.13.X

Markham, Anette N. (2020): Taking data literacy to the streets: Critical pedagogy in the public sphere, in: Qualitative Inquiry 26(2), 227–237. https://doi.org/10.1177/1077800419859024

Menon, Sneha (2017): Aadhaar-based biometric authentication for PDS and food security: Observations on implementation in Jharkhand's Ranchi district, in: Indian Journal of Human Development 11(3), 355–383. https://doi.org/10.1177/0973703017748384

Mhlambi, Sabelo (2020): From rationality to relationality: Ubuntu as an ethical and human rights framework for artificial intelligence governance (Carr Center Discussion Paper Series No. 2020-009), Harvard: Harvard Kennedy School.

Milan, Stefanie/Treré, Emiliano (2019): Big Data from the South(s): Beyond data universalism, in: Television & New Media 20(4), 319–335. https://doi.org/10.1177/1527476419837739

Mohamed, Shakir/Png, Marie-Therese/Isaac, William (2020): Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence, in: Philosophy & Technology 33(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Morung Express (2018, May 16): Denial of food ration due to Aadhaar significant: Report, online at: https://morungexpress.com/denial-food-ration-due-aadhaar-significant-report (last access: 11.26.2025).

Moyo, Last (2020): Decolonial research methodologies: Resistance and liberatory approaches, in: Moyo, Last (Ed.): The decolonial turn in media studies in Africa and the global south (pp. 109–126), Basingstoke: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-52832-4_6

Muldoon, James/Boxi, A Wu (2023): Artificial intelligence in the colonial matrix of power, in: Philosophy & Technology 36(4), Article 88. https://doi.org/10.1007/s13347-023-00687-8

Niesyto, Horst (2017): Medienpädagogik und digitaler Kapitalismus: Für die Stärkung einer gesellschafts- und medienkritischen Perspektive, in: MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung 27, 1–29. https://doi.org/10.21240/mpaed/27/2017.01.13.X

Perrigo, Billy (2022, February 14): Inside Facebook's African sweatshop, in: TIME Magazine, online at: https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/ (last access: 11.26.2025).

Perrigo, Billy (2023, January 18): OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic, in: TIME Magazine, online at: https://time.com/6247678/openai-chatgpt-kenya-workers/ (last access: 11.26.2025).

Petrov, Alexsandar/La Malfa, Emanuele/Torr, Philip/Bibi, Adel (2023): Language model tokenizers introduce unfairness between languages, in: Advances in Neural Information Processing Systems 36, 18768–18782. https://doi.org/10.48550/arXiv.2305.15425

Plantin, Jean-Christophe/Lagoze, Carl/Edwards, Paul N./Sandvig, Christian (2018): Infrastructure studies meet platform studies in the age of Google and Facebook, in: New Media & Society 20(1), 293–310. https://doi.org/10.1177/1461444816661553

Quijano, Aníbal (2000): Coloniality of power and Eurocentrism in Latin America, in: International Sociology 15(2), 215–232. https://doi.org/10.1177/0268580900015002005

Roberts, Sarah T. (2019): Behind the screen: Content moderation in the shadows of social media, Dunmore: Yale University Press.

Scholz, Trebor (2016): Platform cooperativism: Challenging the corporate sharing economy. Rosa Luxemburg Stiftung, online at: https://rosalux.nyc/wp-content/uploads/2020/11/RLS-NYC_platformcoop.pdf (last access: 11.26.2025).

Sesink, Werner (2014): Eine kritische Bildungstheorie der Medien, in: Marotzki, Winfried/Meder, Norbert (Eds.): Perspektiven der Medienbildung (pp. 11–44), Heidelberg: Springer VS. https://doi.org/10.1007/978-3-658-03529-7_2

Sims, Danica A. (2024): How to … do decolonial research, in: The Clinical Teacher 21(5), Article e13806. https://doi.org/10.1111/tct.13806

Spivak, Chakravorty C. (1988): Can the subaltern speak?, in: Nelson, Cary/Grossberg, Lawrence (Eds.): Marxism and the interpretation of culture (pp. 271–313), Champaign: University of Illinois Press.

Srnicek, Nick (2017): Platform capitalism, Boston/New York: Polity Press.

Staab, Philipp (2019): Digitaler Kapitalismus: Markt und Herrschaft in der Ökonomie der Unknappheit, Berlin: Suhrkamp.

Stockwell, Billy (2024, December 21): Facebook inflicted 'lifelong trauma' on content moderators in Kenya, in: CNN, online at: https://edition.cnn.com/2024/12/22/business/facebook-content-moderators-kenya-ptsd-intl (last access: 11.26.2025).

Teklehaymanot, Hailay K./Nejdl, Wolfgang (2025): Tokenization disparities as infrastructure bias: How subword systems create inequities in LLM access and efficiency, in: arXiv. https://doi.org/10.48550/arXiv.2510.12389

Thambinathan, Vivetha/Kinsella, Elizabeth A. (2021): Decolonizing methodologies in qualitative research: Creating spaces for transformative praxis, in: International Journal of Qualitative Methods 20, Article 16094069211014766. https://doi.org/10.1177/16094069211014766

Thatcher, Jim/O'Sullivan, David/Mahmoudi, Dillon (2016): Data colonialism through accumulation by dispossession: New metaphors for daily data. Environment and Planning D, in: Society and Space 34(6), 990–1006. https://doi.org/10.1177/0263775816633195

Tygel, Alan F./Kirsch, Rosana (2016): Contributions of Paulo Freire for a critical data literacy: A popular education approach, in: The Journal of Community Informatics 12(3), 108–121. https://doi.org/10.15353/joci.v12i3.3279

Udah, Hyacinth (2024): Decolonising research for justice: Ethical imperatives and practical applications, in: International Journal of Qualitative Methods 23, Article 16094069241294040. https://doi.org/10.1177/16094069241294040

Udupa, Sahana/Maronikolakis, Antonis/Wisiorek Axel (2023): Ethical scaling for content moderation: Extreme speech and the (in)significance of artificial intelligence, in: Big Data & Society 10(1), 20(1), 74–88. https://doi.org/10.1177/205395172311724

UNHCR (2020): Modernizing registration and identity management in UNHCR: Introducing PRIMES, online at: https://www.unhcr.org/blogs/modernizing-registration-identity-management-unhcr/ (last access: 11.26.2025).

U.S. House of Representatives (2020): Investigation of competition in digital markets: Majority staff report and recommendations (Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary), online at: https://www.govinfo.gov/content/pkg/CPRT-117HPRT47832/pdf/CPRT-117HPRT47832.pdf (last access: 11.26.2025).

Vienna Manifesto on Digital Humanism (2019), online at: https://dighum.ec.tuwien.ac.at/ (last access: 11.26.2025).

Zembylas, Michalinos (2025): Decolonial pathways in education: Walter Mignolo, epistemic delinking, and the risks of ethno-essentialism, in: Globalisation, Societies and Education 1–15. Advance online publication. https://doi.org/10.1080/14767724.2025.2459110

Zuboff, Shoshana (2019): The age of surveillance capitalism: The fight for a human future at the new frontier of power, New York: PublicAffairs.