AI and the Reproduction of Educational Inequalities: A Sociological Perspective

Elena-María García-Alonso
Sociology and Communication Department, University of Salamanca
ORCID: 0000-0003-4902-4225
Email: garelena@usal.es

1. Abstract

This paper examines how Artificial Intelligence (AI) shapes the reproduction of educational inequalities from a sociological perspective. It argues that AI systems are not neutral tools but socio-technical artifacts embedded in power relations that influence knowledge, pedagogical practices, and structures of legitimacy.

The analysis highlights three dimensions through which inequality is reproduced: digital divides that go beyond access and include skills and algorithmic literacy; structural and cultural biases embedded in AI processes such as admission, assessment, and adaptive learning; and algorithmic governmentality, which governs education through surveillance, metrics, and automated feedback, producing new forms of subjectivity oriented toward efficiency rather than critical thinking.

The article concludes that AI in education is an ethical and political challenge. It calls for inclusive and contextualized platform design, critical training for teachers, participatory co-design of evaluation systems, and governance mechanisms that ensure transparency and accountability. From the sociology of education, promoting digital environments that are democratic, inclusive, and emancipatory is essential to achieve educational justice in the age of AI.

Keywords: Educational inequality; Artificial intelligence in education; Digital divide; Algorithmic bias; Algorithmic governance; Algorithmic capital; Critical digital literacy

1. Introducción

In recent years, we have witnessed a growing technologization of education, a process particularly accelerated by the COVID-19 pandemic, which boosted the digitalization of both school and university processes. With the emergence of Artificial Intelligence (AI), the use of adaptive platforms, automated assessments, and virtual learning environments has multiplied. Platforms such as Khan Academy, Duolingo, or Coursera have incorporated conversational agents that mediate learning pathways, while systems such as ChatGPT or Copilot have become popular among students and teachers as tools for consultation, writing assistance, or problem-solving support (García-Alonso et al., 2024). This massive deployment not only highlights the pedagogical potential of AI, but also the risks of unequal adoption: while those with greater technological resources manage to integrate these tools in creative ways, precarious contexts often struggle to meet even the most basic connectivity needs.

There is broad consensus on the effectiveness of these tools in improving efficiency and facilitating access to content (García-Alonso et al., 2024; Hakiki et al., 2023; Zain Abdillah et al., 2023). The spread of conversational chatbots has marked a turning point in the educational field (Holmes & Tuomi, 2022; Zawacki-Richter, 2024). These technologies open up significant possibilities: from personalized learning support to the automation of teaching and administrative tasks. However, international evidence shows that even as basic access improves, material inequalities persist (devices, connectivity) and, more importantly, inequalities of use and skills: performance in Computer and Information Literacy (CIL) and Computational Thinking (CT) varies by socioeconomic status and territory; a considerable proportion of students lack basic digital skills; and teachers frequently report insufficient preparation to critically integrate technology (IEA, 2023; van de Werfhorst et al., 2022; Zawacki-Richter, 2024) . The result of this unequal integration is an intensification of educational divides, not only in terms of access (the first digital divide), but also in the meaningful use of technology (the second digital divide), and in the ability to understand and critically interrogate the algorithmic systems mediating learning (algorithmic capital) (Van Dijk, 2017).

We are thus confronted with a phenomenon that goes beyond technology, transforming relationships, roles, and opportunities within the educational system (Selwyn, 2022). The algorithms underpinning Artificial Intelligence tend to reproduce and amplify pre-existing inequalities, configuring new forms of exclusion linked to algorithmic capital (Couldry & Mejias, 2020; Williamson et al., 2023). In this sense, we are entering a scenario of growing algorithmic governance, in which decision-making and regulation increasingly shift toward opaque systems that shape educational opportunities and trajectories (Amoore, 2020; Beer, 2017).

From the perspective of the sociology of education, the study of how schools contribute to the reproduction of inequality is a well-established field of research. Sociologists such as Bourdieu, Foucault, and Giddens have shown, from different approaches, how education not only transmits knowledge but also legitimizes and reproduces the status quo. Today, with the irruption of Artificial Intelligence (AI) and the profound changes in how we engage with learning and technology, it becomes necessary to reinterpret these classical theories in light of the new algorithmic context. Contemporary debates on digital divides, algorithmic capital, and digital governance expand this framework of understanding while maintaining dialogue with the critical roots of the sociology of education.

Concerns about inequality and the challenges arising from the integration of AI in education are also present in ethical and regulatory debates, where emphasis is placed on the need to incorporate principles of ethics and social justice in AI implementation. UNESCO (UNESCO, 2021) highlights the urgency of ensuring a human rights–centered approach that avoids reproducing biases and exclusions, while the OECD (*OECD Digital Education Outlook 2023*, 2023) has underscored the importance of training teachers in critical algorithmic literacy, and the European Union (European Commission, n.d.) has launched regulatory initiatives aimed at ensuring transparency and accountability in automated systems. This global framework reinforces the relevance of sociological analysis by positioning educational AI not as a mere technical advancement, but as a political and cultural phenomenon that redefines the relationship between knowledge, power, and inequality.

This paper is conceived as a theoretical review aimed at analyzing the transformations introduced by Artificial Intelligence (AI) in the educational field from a sociological perspective of inequalities. It is not a systematic review in the strict sense, but rather an exercise in conceptual analysis that seeks to articulate different theoretical traditions to understand new forms of social reproduction in a context of increasing datafication and automation of educational processes. This raises a key question: how can classical theories of educational inequality be reinterpreted in the age of algorithms and artificial intelligence?

Digital technologies, and algorithmic systems in particular, should be understood not as mere technical instruments, but as sociotechnical artifacts embedded in power relations and historical dynamics of exclusion (Selwyn & Jandrić, 2020). This starting point makes it possible to articulate an analytical framework that combines contributions from the sociology of education, theories of social reproduction, research on the digital divide, and critical perspectives in science and technology studies (STS), with particular attention to the notion of algorithmic governmentality. On this basis, the chapter proposes interpreting algorithmic educational inequality through three interrelated dimensions: digital divides and algorithmic capital, social reproduction and structural biases, and algorithmic governmentality as a form of socialization. Together, these dimensions configure a web of exclusions that operate simultaneously, mutually reinforcing one another within the school experience mediated by intelligent technologies.

The methodology used in this study relies on a critical review designed to recover the main theoretical lines that explain educational inequality from a sociological perspective applied to the development of AI. To this end, a search strategy was carried out in academic databases, focusing on the period 2022–2025, coinciding with the generalization of conversational chatbots. The review prioritized recent articles and monographs, as well as works grounded in classical theoretical frameworks (Bourdieu, Foucault, Giddens). The final selection combined empirical and theoretical works, with special emphasis on those explicitly addressing the reproduction of inequalities and digital governance in educational contexts, thereby ensuring both conceptual coherence and analytical relevance.

2. Sociological Theory in the Face of the Emergence of Artificial Intelligence in Education: Between Digital Inequality and Algorithmic Governmentality

The following section presents three dimensions identified as key elements for understanding how AI influences processes of social inequality. These dimensions do not operate independently; rather, their effects intersect and mutually reinforce each other, shaping complex shared spaces that allow for multiple theoretical readings.

2.1 Algorithmic Capital and Digital Divides

Bourdieu (Bourdieu & Passeron, 1990) introduced the notion of cultural capital as a strategic resource that decisively shapes the reproduction of educational inequalities. For this author, the language and forms of knowledge of dominant groups are presented as legitimate and universal. In this way, cultural capital becomes an invisible criterion of distinction which, together with school evaluation mechanisms and the sustaining habitus, consolidates persistent social hierarchies. Those who possess greater cultural capital see their position reinforced, while those who lack it are relegated.

In the current context, this cultural capital is projected and reconfigured as digital cultural capital, understood as the set of skills, knowledge, and dispositions that enable meaningful and reflective participation in digital environments (Sefton-Green et al., 2016). Digital cultural capital operates similarly to the cultural capital defined by Bourdieu, that is, as yet another mechanism of social reproduction. The competencies required by automated teaching systems, such as self-directed autonomy, formal argumentative competence, or the ability to manage multiple platforms, mirror and reinforce dominant habitus. AI technologies privilege certain student profiles, rendering other more situated or contextual forms of knowledge invisible, thereby reinforcing pre-existing educational hierarchies (Katiyar et al., 2024; Madianou, 2021). This type of capital acts as a symbolic resource and facilitates the reproduction of educational inequalities through improved academic performance and more active participation in the design and use of technologies (Celik, 2023).

When delving into the functioning, language, and logic of AI systems, the accumulation of digital capital is transformed, in more technical terms, into a new form of algorithmic capital, understood as the capacity to comprehend, influence, or intervene in the operating logics of automated systems that structure the educational experience (Beer, 2017). We are therefore confronted with new forms of technological capital that interact within the educational arena: digital capital and algorithmic capital. Digital capital constitutes a necessary but insufficient prerequisite for navigating AI-mediated educational environments. The notion of algorithmic capital introduces an additional level of distinction. For example, while one group of students may limit themselves to using a chatbot as a superficial consultation tool, others are able to formulate advanced prompts, critically evaluate responses, and reorient interactions to achieve deeper learning. This difference, seemingly technical, translates into unequal access to learning opportunities: those who possess algorithmic capital are able to appropriate the tool as a means of intellectual autonomy, while those who lack it remain in a more passive and dependent position (Zhai et al., 2024).

The unequal accumulation of these capitals translates into concrete forms of exclusion, manifested in what are commonly referred to as digital divides. Students with greater cultural capital and familiarity with digitized school codes benefit from these environments, whereas those from contexts with lower technological exposure face not only material barriers (the first digital divide) but also symbolic and practical ones (the second digital divide), which can be understood in terms of access, capacity for use, appropriation, and impact (Jan A.G.M. van Dijk, 2012).

The first digital divide refers to inequalities in access to basic technological infrastructure, devices, connectivity, and platforms, a necessary but insufficient condition for the effective appropriation of AI in schools. Various studies document how the introduction of algorithmic learning systems makes these inequalities more visible and severe, restricting the benefits of such technologies to privileged schools (Area-Moreira et al., 2024; Kharchenko et al., 2024). This situation is especially critical in rural contexts (Castro et al., 2025; Villarino, 2025) and in the Global South (Regmi, 2024), where the lack of specific policies and investment in digital infrastructure perpetuates historical exclusions and reinforces the global hierarchy of knowledge. Even in high-connectivity contexts, such as that analyzed by Zhang et al. (Zhang et al., 2025) in the United States, significant inequalities persist in the acceptance and use of AI according to demographic and social variables: students from lower socioeconomic strata and racialized backgrounds make less creative and effective use of ChatGPT than their peers.

These inequalities become even more apparent when comparing global contexts. According to data from the International Telecommunication Union (International Telecommunication Union (ITU) & UNESCO Broadband Commission, 2022), more than 90% of European households have access to fixed broadband internet, while in Africa the figure barely exceeds 40%. The divide is not only geographical but also socioeconomic: in many rural areas of Latin America, fewer than 50% of households have stable connections, significantly limiting the educational use of AI. This inequality in infrastructure translates into a global hierarchy of knowledge, in which data generated in privileged contexts feed algorithmic systems, while countries with lower connectivity remain merely passive consumers of tools developed in the Global North (Valencia-Londoño et al., 2025)

The second digital divide, in turn, refers to differences in the capacity for meaningful, autonomous, and critical use of technologies. This dimension is particularly relevant in the case of AI, given the opaque and complex nature of its operating systems. Thus, the absence of computational thinking and algorithmic literacy limits the possibilities for active intervention in digital processes by both teachers and students (Bian et al., 2024; Celik, 2023; Sharma & Gupta, 2024). A large proportion of practicing teachers lack specific training in digital literacy and, even more so, in algorithmic literacy. As a result, although they may be able to use AI-based educational tools, they depend on them without possessing critical criteria to interpret recommendations or detect biases in their outputs. In this way, the lack of algorithmic capital among teachers limits their professional agency and reinforces dependence on automated systems, reducing their capacity to adapt pedagogical decisions to the concrete contexts of the classroom (Agca & Korkmaz, 2025; Arantes, 2023; Castro et al., 2025). This divide also tends to reproduce itself generationally, as students are socialized in educational models where teacher judgment is subordinated to algorithmic criteria (Mac Fadden et al., 2024).

2.2 Reproduction of Inequalities and Algorithmic Biases

In sociology, the reproduction of inequalities has been analyzed through theoretical frameworks that reveal how social structures legitimize and perpetuate dominant positions without the need for explicit coercion. Bourdieu (1990) emphasized the role of habitus and cultural capital in the intergenerational transmission of privilege, showing that schools tend to value the cultural codes of the elites while delegitimizing other forms of knowledge. From another perspective, Foucault (1980) highlighted the disciplinary and normalizing dimensions of modern institutions, which, through devices of control, shape subjects according to standards of performance and conduct. Giddens (2012) proposed the duality of structure, whereby everyday practices reproduce the very social conditions that enable them, demonstrating that inequality is not accidental but a recurring effect of both action and structure.

In this regard, these issues can also be analyzed through the theory of "invisible infrastructures" proposed by Bowker and Star (Bowker & Star, 2000), who argue that technical systems crystallize social and political norms that are rarely visible to their users. In education, algorithms do not merely classify data; they embed cultural assumptions and criteria of legitimacy within their designs that largely go unnoticed. For instance, algorithms favor students who possess higher levels of cultural and digital capital, that is, competencies in technological literacy, computational thinking, and self-regulated learning (Zhang et al., 2025), while marginalizing alternative options.

Artificial Intelligence relies on large volumes of historical data processed through machine learning algorithms. These data reflect the social, cultural, and economic contexts in which they were generated. In other words, although in theory AI should function as a neutral and impartial entity, in practice it operates on historically biased data and complex technical criteria (algorithms) that are rarely problematized. This contributes to the reproduction of structural, symbolic, and cultural inequalities within the educational sphere, especially when these technologies are deployed without considering the social and institutional conditions in which they are embedded (Ricoy-Casas et al., 2025; Roshanaei, 2024; Swist & Gulson, 2023). Within this framework, digital discrimination derived from systems designed for an "ideal" user, generally male, urban, and from privileged sectors (Beloeva & Venelinova, 2024; Francesca Gottschalk & Crystal Weise, 2023), exacerbates the exclusion of vulnerable groups. Such design biases produce what Gan and Bai (Gan & Bai, 2023) term *epistemic exclusion*: the invisibilization of subaltern knowledges in the face of dominant pedagogical models imposed by algorithmic logics.

Recent studies on the application of Artificial Intelligence (AI) in education warn that the reproduction of inequalities is not an accidental side effect but a structural risk inherent to the design, implementation, and operating logic of algorithmic systems. The reproduction of inequality is particularly evident in admission and school choice processes, where algorithms, lacking ethical audits and transparency, reinforce dynamics of racial and class segregation (Swist & Gulson, 2023).

This logic also persists in higher education, where automated admission systems tend to exclude underrepresented groups because they rely on historical data that reflect structural biases (Roshanaei, 2024). In the United Kingdom, the algorithm used in 2020 to predict university entrance grades systematically penalized students from public schools compared to those from private institutions, reproducing class and territorial inequalities (Saqib Safdar, 2025). In the United States, the COMPAS case revealed how algorithmic prediction systems in criminal justice reinforced racial bias under the guise of neutrality (Dressel & Farid, 2018). During the pandemic, the use of surveillance algorithms such as online exam proctoring was reported to discriminate against racialized students, women, and persons with disabilities, evidencing how algorithmic bias translates into practical exclusion and experiences of educational injustice (Selwyn & Jandrić, 2020). In response, the literature emphasizes the need for critical management of these tools based on ethical principles and equity frameworks that not only monitor their technical functioning but also incorporate pedagogical and social justice criteria capable of counteracting the reproduction of inequalities and opening spaces for more inclusive forms of knowledge (Mac Fadden et al., 2024).

Algorithms, trained on data reflecting existing social hierarchies, reinforce these logics through hidden processes of classification, prediction, and evaluation, often detached from pedagogical judgment or the specific contexts of individual students (Akgun & Greenhow, 2022). These mechanisms show how institutions legitimize inequality by rewarding dominant capitals, paving the way for further exclusions in domains such as gender. Adaptive tutoring and recommendation systems in STEM environments tend to reproduce gender stereotypes, encouraging male students toward technical trajectories while minimizing female participation in fields such as programming or engineering. These biases do not stem from explicit intentionality, but rather from the historical data that feed the algorithms and from cultural imaginaries that code which profiles are considered appropriate for each field of knowledge (Due et al., 2024). In this way, AI not only reinforces social and racial inequalities but also perpetuates traditional gender roles, restricting the educational and professional opportunities of women and other underrepresented gender identities in scientific and technological fields. At the same time, it strengthens logics of symbolic exclusion, whereby the styles, codes, and aesthetics associated with racialized youth cultures or community practices are systematically marginalized (Gan & Bai, 2023). The result is the consolidation of a cultural canon that privileges certain forms of knowledge and expression while excluding minority ones.

The functioning of educational AI also produces performative effects on pedagogical practices, defining what is considered legitimate learning, adequate performance, or acceptable conduct. Thus, the automation of assessment and monitoring processes contributes to disciplining and homogenizing educational trajectories, invisibilizing the cultural, social, and cognitive diversity of students (Nopas, 2025; Sharma & Gupta, 2024; Zhang et al., 2024). Along these lines, Sætra (Sætra, 2023) warns that the uncritical integration of generative AI may erode the emancipatory potential of education, consolidating the status quo rather than promoting social transformation.

The effects of AI go beyond reproducing the discourse of dominant groups. Rather than merely processing and reproducing information through algorithms, AI systems deployed in educational settings act as mechanisms of permanent surveillance, where every student interaction, from connection times to quiz responses and navigation within platforms, is recorded and analyzed to feed predictive mechanisms and automated feedback loops. This “datafication” of the learning process imposes standards of conduct and performance that homogenize the diversity of trajectories and contexts, while shaping behaviors and participation styles subject to regulation by the algorithm’s technical parameters (Nopas, 2025; Sharma & Gupta, 2024; Zhang et al., 2024). Such technical normalization constitutes what Zuboff (Zuboff, 2019) called a “pedagogy of control,” in which student subjectivity is continually adjusted to imperatives of efficiency and optimization, reducing the agency of both teachers and learners under the appearance of system objectivity.

Finally, by disregarding contextual conditions of use, AI contributes to reinforcing colonial and exclusionary dynamics, particularly at the margins of the global educational system. From a decolonial perspective, Valencia (Valencia-Londoño et al., 2025) argues that the lack of contextualization and epistemic justice in the design and application of these technologies perpetuates forms of educational exclusion. This problem is especially acute in the Global South, where infrastructure and digital literacy limitations exacerbate the exclusionary effects of AI (Mbambo & du Plessis, 2025; Oyelere & Aruleba, 2025; Regmi, 2024). This surveillance regime is embedded within the dynamics of data colonialism, as technological infrastructures and the vast volumes of information generated in the Global South are extracted, processed, and monetized by platforms largely designed in the North (Valencia-Londoño et al., 2025). The extractive logic not only displaces situated knowledges and community pedagogies but also subordinates local educational practices to metrics and algorithmic models conceived without the participation of those most affected. In this way, asymmetries in data governance reproduce colonial power relations, eroding epistemic justice and cultural diversity in the classroom while perpetuating structural socioeconomic inequalities that go far beyond the mere technological access gap (Mbambo & du Plessis, 2025; Oyelere & Aruleba, 2025). Some experiences in Latin America point to alternatives, reconfiguring datasets to include local knowledges and minority languages, thereby countering data colonialism with inclusive approaches that value cultural diversity (Valencia-Londoño et al., 2025).

2.3 Algorithmic Governmentality as an Agent of Educational Socialization

The Foucauldian notion of governmentality provides a fundamental framework for understanding contemporary forms of the exercise of power. Foucault (Foucault & Gordon, 1980) emphasized rationalities that operate beyond formal laws and classical institutions, unfolding through devices, norms, and techniques that shape everyday life. To govern does not simply mean to surveil or control from the outside, but to "conduct conduct," guiding the way individuals relate to themselves and to others. Power, in this sense, is not imposed exclusively in a repressive manner; it acts by shaping subjectivities, molding how individuals think, decide, and act.

In the case of Artificial Intelligence applied to education, this "conduct of conduct" materializes paradigmatically. Algorithmic systems that mediate the school experience induce students to self-regulate according to predefined indicators: progress bars, automated feedback, pathway suggestions, or alerts about expected performance. It is no longer necessary to impose discipline externally, the very technical architecture of the platform stimulates a process of constant self-adjustment, presented as objective and neutral. This is precisely the core of what can be called *algorithmic governmentality*: a form of power exercised through calculation, classification, and prediction, without the mediation of explicit pedagogical deliberation.

Rouvroy y Berns (Rouvroy & Berns, 2013) introduced the notion of algorithmic governmentality to describe the capacity of algorithms to predict and classify behaviors, orienting social trajectories without direct intervention from traditional institutions. Translated into the educational domain, this means that key decisions regarding what is learned, how performance is assessed, or which trajectories are deemed desirable become conditioned by technical criteria inscribed in automated systems. Thus, processes of measurement, continuous feedback, and prediction progressively displace pedagogical judgment in favor of standards of efficiency, traceability, and optimization.

Algorithmic governmentality also has a socializing function. It does not merely transmit information but teaches individuals how to behave, organize their time, and conceive of their academic progress. It is a mechanism of subject formation that turns the learning experience into a practice guided by invisible technical norms.

Algorithmic systems deploy various strategies to exercise their influence on the learning experience. First, recommendation engines and adaptive platforms (for example, YouTube EDU or personalized learning software) select and prioritize content according to metrics of performance, popularity, or user profile suitability. In this way, they determine which forms of knowledge are made visible and which are discarded, restricting exposure to critical or alternative perspectives and homogenizing the learning horizon (Ricoy-Casas et al., 2025).

By establishing metrics, thresholds, and classifications, AI systems configure new rules of the educational game. Automated assessment is no longer limited to measuring learning; it produces categories, legitimizes trajectories, and makes decisions about educational success or failure (Selwyn, 2022). Digital environments, by recording every interaction, act as permanent surveillance devices that feed into mechanisms of automated feedback. Lyon (Lyon, 2018) describes this phenomenon as a form of pedagogy of self-adjustment: students become accustomed to regulating their behavior according to invisible indicators, internalizing performance norms. Zuboff (Zuboff, 2019), in this vein, speaks of a pedagogy of control, in which student subjectivity is continually adjusted to imperatives of efficiency and optimization.

The issue does not concern students alone. Teachers also experience a profound transformation. Giddens (Giddens, 2012) noted that contemporary societies depend on expert systems inaccessible to most. When these “experts” are algorithms, the opacity of their operational rules erodes teachers’ professional agency, limiting the incorporation of contextual criteria or the exercise of ethical-pedagogical judgment. This erosion materializes in the outsourcing of functions traditionally associated with teachers: evaluating, recommending, and determining learning pathways. As Arantes (Arantes, 2023), indicates, AI-mediated education tends to prioritize quantitative criteria, displacing reflective practices.

Concrete examples illustrate this loss of agency. In some Learning Management Systems (LMS), grades are generated automatically from standardized rubrics, reducing the teacher’s role to supervising results previously calculated by the machine. Other AI platforms design personalized pathways that recommend content, exercises, and work rhythms without the need for teacher mediation. In this framework, the teacher's role shifts toward administrative management or technical support, while central pedagogical decisions are externalized. This limits the incorporation of situated knowledge and professional expertise into educational processes and weakens the capacity to adapt to the particularities of each group.

At the same time, the figure of the *algorithmic subject* emerges: an individual whose learning orientation is structured through indicators such as progress bars, badges, or exercise suggestions that quantify the school experience (Ricoy-Casas et al., 2025; Selwyn, 2022). Gamification, widely extended in digital platforms, turns learning into a quantified practice: value lies not in critical reflection or creativity but in the achievement of objective, standardized goals. In this process, qualitative aspects of education, such as collaboration, imagination, or critical thinking, tend to be relegated. The algorithmic subject responds to the logic of the platform, shaping ways of thinking and acting aligned with principles of efficiency, self-optimization, and technical control.

Algorithmic governmentality is not an inescapable condition. Alternatives exist that seek to rebalance power within the classroom. One of these is critical digital literacy, which aims to equip students with tools to denaturalize the ideologies embedded in digital platforms. Gan and Bai (Gan & Bai, 2023) argue that, from Marxist and postcolonial perspectives, students can be trained to recognize how algorithms privilege certain knowledges while marginalizing others.

Another line of resistance lies in teacher-led audits, which confront the results of automated evaluation with professional judgment, thereby reclaiming space for pedagogical deliberation (Sharma & Gupta, 2024). Such practices assert teachers' agency and highlight the need to complement technical assessments with ethical and contextual criteria. In parallel, participatory co-design experiences have emerged in which students and teachers collaborate in defining metrics and evaluation systems, avoiding the unilateral imposition of algorithmic criteria.

These initiatives demonstrate that algorithmic governmentality is not a closed destiny but a political and pedagogical space under dispute. Critical socialization, which implies learning not only *through* AI but also *about* AI, emerges as a fundamental strategy to preserve educational agency. Only through such practices is it possible to counter the hegemony of logics of efficiency and control and to construct a more democratic, inclusive, and emancipatory model of digital education (Regmi, 2024; Ricoy-Casas et al., 2025)

3. Conclusions

The analysis developed shows how the integration of Artificial Intelligence in education operates simultaneously along three interrelated axes: digital divides and algorithmic capital, social reproduction and structural biases, and algorithmic governmentality as a form of socialization. These dimensions do not act in isolation but are articulated in a systemic process of production and reproduction of inequalities: divides in access to and use of technologies condition which data feed the systems; these data reinforce social hierarchies through algorithmic biases; and these hierarchies are consolidated through dynamics of surveillance, classification, and self-adjustment characteristic of algorithmic governmentality. In this way, exclusionary mechanisms feed into each other, shaping a new educational regime marked by datafication and automation.

From the field of the sociology of education, this work offers a novel articulation between Bourdieu's studies on cultural capital, the Foucauldian perspective of governmentality, and decolonial and Science, Technology and Society (STS) approaches. By integrating these theoretical frameworks, it becomes evident that AI systems are not mere technological instruments but artifacts loaded with power relations that negotiate knowledge, pedagogical practices, and structures of legitimacy. In this sense, educational AI emerges as a space of political and symbolic struggle, in which the contours of epistemic justice and professional agency are defined.

The results of this study suggest several lines of intervention to mitigate the identified risks and to promote a fairer AI. First, platform design must be inclusive and contextualized, actively incorporating local knowledges and minority languages into datasets in order to reduce epistemic exclusion. Second, critical algorithmic literacy training for teachers is essential, enabling them to interpret and question automated recommendations and to complement technical judgments with ethical pedagogical criteria. Moreover, schools should establish data ethics committees and subject their systems to regular audits that guarantee algorithmic transparency and allow school communities to participate in governance processes. Finally, co-design policies that involve both students and teachers in defining performance metrics can prevent the imposition of homogeneous indicators and foster criteria of educational justice.

Some countries and international organizations have already begun to implement measures in this direction. UNESCO (UNESCO, 2021) has developed a global reference framework emphasizing the need for human rights–centered AI, with a focus on equity, inclusion, and cultural diversity. In Finland (Tedre et al., 2023), participatory co-design programs have been developed in which teachers, students, and technology designers collaborate in the creation of educational tools, ensuring that performance metrics respond to pedagogical rather than purely technical criteria. In Canada, several school districts have promoted community governance initiatives for educational data, involving families and local communities in decision-making. By contrast, the case of China (Andrejevic & Selwyn, 2020) illustrates how the massive implementation of AI technologies in schools can result in dynamics of social control, where facial surveillance and attention monitoring systems reinforce a disciplinary pedagogy oriented toward obedience rather than emancipation. These examples demonstrate that the regulation and use of AI in education are not neutral but respond to political and cultural choices that determine their social effects.

This conceptual exercise has certain limitations. First, it lacks direct empirical evaluation: the absence of case studies or field data prevents the punctual validation of the theoretical dynamics proposed. Likewise, it does not delve into the affective and relational dimension of educational AI, that is, the way in which these technologies shape emotions, motivations, and interpersonal bonds within the classroom.

To advance the understanding of these phenomena, several research paths are proposed. First, it is necessary to design mixed empirical studies that combine content analysis of educational platforms with classroom ethnographies and surveys of teachers and students about their experiences with AI. Second, longitudinal studies could reveal how algorithmic capital evolves across different social groups and what impact it has on academic trajectories. Third, participatory co-design experiments would make it possible to evaluate the effectiveness of dashboards and recommendation systems developed jointly with school communities. Finally, interdisciplinary research that incorporates educational psychology and STS studies will be essential to explore the affective and relational dimensions of algorithmic governmentality.

In the age of Artificial Intelligence, educational justice transcends the mere guarantee of access to technology and requires an active confrontation with the logics of power embedded in algorithms. Only through critical analysis, democratic participation, and the co-design of tools can we build an educational AI model that celebrates cultural diversity, strengthens teacher agency, and promotes the emancipation of all those involved in the learning process.

Ultimately, the question of Artificial Intelligence in education cannot be addressed solely as a technical challenge linked to system efficiency, but as an ethical, political, and pedagogical challenge of global scope. What is at stake is not only how algorithms are integrated into classrooms, but also which conception of justice, which model of knowledge, and which forms of coexistence are privileged in the process. Approaching AI from this perspective implies recognizing it as a contested space of possibilities, where education can either reproduce existing hierarchies or become a laboratory of inclusive and emancipatory practices. This, ultimately, is the horizon that this book invites us to explore: the need to think of educational AI not as an inevitable destiny, but as a field open to critique, imagination, and democratic transformation.

References

Agca, R. K., & Korkmaz, O. (2025). Experimental Perspective on Artificial Intelligence Anxiety. *International Journal of Technology in Education*, *8*(1), 22–44. https://doi.org/10.46328/ijte.846

AI and education: guidance for policy-makers. (2021). UNESCO. https://doi.org/10.54675/PCSP7350

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Amoore, L. (2020). *Cloud Ethics*. Duke University Press. https://doi.org/10.1215/9781478009276

Andrejevic, M., & Selwyn, N. (2020). Facial recognition technology in schools: critical questions and concerns. *Learning, Media and Technology*, *45*(2), 115–128. https://doi.org/10.1080/17439884.2020.1686014

Arantes, J. A. (2023). Personalization in Australian K-12 classrooms: how might digital teaching and learning tools produce intangible consequences for teachers' workplace conditions? *Australian Educational Researcher*, *50*(3), 863–880. https://doi.org/10.1007/s13384-022-00530-7

Area-Moreira, M., Del Prete, A., Sanabria-Mesa, A. L., & Sannicolás-Santos, M. B. (2024). No todas las herramientas de IA son iguales. Análisis de aplicaciones inteligentes para la enseñanza universitaria. *Digital Education Review*, *45*, 141–149. https://doi.org/10.1344/der.2024.45.141-149

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, *20*(1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147

Beloeva, S., & Venelinova, N. (2024). Digital Discrimination Risks in the Transformation of Higher Education. *Strategies for Policy in Science and Education-Strategii Na Obrazovatelnata i Nauchnata Politika*, *32*(5s), 98–108. https://doi.org/10.53656/str2024-5s-9-dig

Bian, Q., Ling, X., & Yan, S. (2024). Bridging the Algorithmic Divide. *Education as Change*, *28*. https://doi.org/10.25159/1947-9417/17983

Bourdieu, P. ., & Passeron, J.-C. . (1990). *Reproduction in education, society and culture. Revised edition*. Sage Publications.

Bowker, G. C. ., & Star, S. Leigh. (2000). *Sorting things out : classification and its consequences*. MIT Press.

Castro, A., Díaz, B., Aguilera, C., Prat, M., & Chávez-Herting, D. (2025). Identifying Rural Elementary Teachers' Perception Challenges and Opportunities in Integrating Artificial Intelligence in Teaching Practices. *Sustainability (Switzerland)*, *17*(6). https://doi.org/10.3390/su17062748

Celik, I. (2023). Exploring the Determinants of Artificial Intelligence (AI) Literacy: Digital Divide, Computational Thinking, Cognitive Absorption. *Telematics and Informatics*, *83*. https://doi.org/10.1016/j.tele.2023.102026

Couldry, N., & Mejias, U. A. (2020). *The Costs of Connection*. Stanford University Press. https://doi.org/10.1515/9781503609754

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, *4*(1). https://doi.org/10.1126/sciadv.aao5580

Due, S., Das, S., Andersen, M., López, B. P., Nexø, S. A., & Clemmensen, L. (2024). *Evaluation of Large Language Models: STEM education and Gender Stereotypes*.

European Commission. (2025). *Tools for schools and educators Digital Education Action Plan: policy background*.

Foucault, Michel., & Gordon, Colin. (1980). *Power/knowledge : selected interviews and other writings 1972-1977*. Pantheon Books.

Francesca Gottschalk, & Crystal Weise. (2023). *Digital equity and inclusion in education*. https://doi.org/10.1787/7cb15030-en

Gan, F., & Bai, Q. (2023). Karl Marx's thoughts on critical pedagogy, reproduction, and aesthetic literacy in STEAM education and praxis. *Educational Philosophy and Theory*, *55*(13), 1513–1525. https://doi.org/10.1080/00131857.2023.2197208

García-Alonso, E. M., León-Mejía, A. C., Sánchez-Cabrero, R., & Guzmán-Ordaz, R. (2024). Training and Technology Acceptance of ChatGPT in University Students of Social Sciences: A Netcoincidental Analysis. *Behavioral Sciences*, *14*(7), 612. https://doi.org/10.3390/bs14070612

Giddens, Anthony. (2012). *The consequences of modernity*. Polity.

Hakiki, M., Fadli, R., Samala, A. D., Fricticarani, A., Dayurni, P., Rahmadani, K., Astiti, A. D., & Sabir, A. (2023). Exploring the impact of using Chat-GPT on student learning outcomes in technology learning: The comprehensive experiment. *Advances in Mobile Learning Educational Research*, *3*(2), 859–872. https://doi.org/10.25082/AMLER.2023.02.013

Holmes, W., & Tuomi, I. (2022). State of the art and practice in <scp>AI</scp> in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

IEA. (2023). *ICILS 2023 International Report, Results, and Resources Now Available Online*.

International Telecommunication Union (ITU) & UNESCO Broadband Commission. (2022). *The State of Broadband 2022: Accelerating Broadband for New Realities*. ITU.

Jan A.G.M. van Dijk. (2012). *The Evolution of the Digital Divide - The Digital Divide turns to Inequality of Skills and Usage*. IOS Press.

Katiyar, Prof. Dr. N., Awasthi, Mr. V. K., Pratap, Dr. R., Mishra, Mr. K., Shukla, Mr. N., Singh, Mr. R., & Tiwari, Dr. M. (2024). Ai-Driven Personalized Learning Systems: Enhancing Educational Effectiveness. *Educational Administration Theory and Practices*. https://doi.org/10.53555/kuey.v30i5.4961

Kharchenko, A., Nalyvaiko, O., Kreydun, N., Sheiko, A., Ptushka, A., Khatuntseva, S., & Zotova, L. (2024). Digital Technologies as a Factor of Transformation of Learning in the University Education. *Revista Romaneasca Pentru Educatie Multidimensionala*, *16*(4), 97–126. https://doi.org/10.18662/rrem/16.4/909

Lyon, David. (2018). *The culture of surveillance : watching as a way of life*. Polity Press.

Mac Fadden, I., García-Alonso, E.-M., & López Meneses, E. (2024). Science Mapping of AI as an Educational Tool Exploring Digital Inequalities: A Sociological Perspective. *Multimodal Technologies and Interaction*, *8*(12), 106. https://doi.org/10.3390/mti8120106

Madianou, M. (2021). Nonhuman humanitarianism: when "AI for good" can be harmful. *Information, Communication & Society*, *24*(6), 850–868. https://doi.org/10.1080/1369118X.2021.1909100

Mbambo, G. P., & du Plessis, E. C. (2025). Evaluating technical vocational education and training college student's digital skills versus throughput rate. *Discover Education*, *4*(1), 6. https://doi.org/10.1007/s44217-025-00396-8

Nopas, D. (2025). Algorithmic learning or learner autonomy? Rethinking AI's role in digital education. *Qualitative Research Journal*. https://doi.org/10.1108/QRJ-11-2024-0282

OECD Digital Education Outlook 2023. (2023). OECD. https://doi.org/10.1787/c74f03de-en

Oyelere, S. S., & Aruleba, K. (2025). A comparative study of student perceptions on generative AI in programming education across Sub-Saharan Africa. *Computers and Education Open*, *8*. https://doi.org/10.1016/j.caeo.2025.100245

Regmi, K. D. (2024). The rise of learning technology in an unequal world: potentials and limitations in enhancing lifelong learning. *International Review of Education*, *70*(3), 433–452. https://doi.org/10.1007/s11159-023-10058-2

Ricoy-Casas, R. M., Fernández-González, R., & Santos-Garrido, M. (2025). Underrepresented students and artificial intelligence | Alumnado inteligencia artificial. *European Public and Social Innovation Review*, *10*. https://doi.org/10.31637/epsir-2025-843

Roshanaei, M. (2024). Towards best practices for mitigating artificial intelligence implicit bias in shaping diversity, inclusion and equity in higher education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12605-2

Rouvroy, A., & Berns, T. (2013). Gouvernementalité algorithmique et perspectives d'émancipation. *Réseaux*, *n° 177*(1), 163–196. https://doi.org/10.3917/res.177.0163

Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society*, *75*. https://doi.org/10.1016/j.techsoc.2023.102372

Saqib Safdar. (2025, August 11). *How an Algorithm Deepened Inequality in UK A-Level Results—and What It Teaches Us About AI Fairness*. Https://Www.Saqibsafdar.Com/Bias-in-the-2020-a-Level-Grading-Algorithm/.

Sefton-Green, J., Marsh, J., Erstad, O., & Flewitt, R. (2016). *Establishing a research agenda for the digital literacy practices of young children: A White Paper for COST Action IS1410*.

Selwyn, N. (2022). The future of span style="font-variant:small-caps;"AI/span and education: Some cautionary notes. *European Journal of Education*, *57*(4), 620–631. https://doi.org/10.1111/ejed.12532

Selwyn, N., & Jandrić, P. (2020). Postdigital Living in the Age of Covid-19: Unsettling What We See as Possible. *Postdigital Science and Education*, *2*(3), 989–1005. https://doi.org/10.1007/s42438-020-00166-9

Sharma, M., & Gupta, S. (2024). *The Role of ChatGPT to Improve Teaching and Learning in Higher Education* (pp. 87–98). https://doi.org/10.4018/979-8-3693-1310-7.ch006

Swist, T., & Gulson, K. N. (2023). School Choice Algorithms: Data Infrastructures, Automation, and Inequality. *Postdigital Science and Education*, *5*(1), 152–170. https://doi.org/10.1007/s42438-022-00334-z

Tedre, M., Mäkitalo, K., Vartiainen, H., Kahila, J., Laru, J., & Iwata, M. (2023). Generation AI: Participatory Machine Learning Co-Design Projects with K-9 Students in Finland. *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 2*, 657–657. https://doi.org/10.1145/3587103.3594171

UNESCO. (2021). *Reimagining our futures together: a new social contract for education*. Educational and Cultural Organization of the United Nations.

Valencia-Londoño, P. A., Cardona-Rodas, H., & Jiménez-Builes, J. A. (2025). Artificial intelligence in education: advancing educational digital inclusion for adults older with diverse neuromuscular conditions. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1464030

van de Werfhorst, H. G., Kessenich, E., & Geven, S. (2022). The digital divide in online education: Inequality in digital readiness of students and schools. *Computers and Education Open*, *3*, 100100. https://doi.org/10.1016/j.caeo.2022.100100

Van Dijk, J. A. G. M. (2017). Digital Divide: Impact of Access. In *The International Encyclopedia of Media Effects* (pp. 1–11). Wiley. https://doi.org/10.1002/9781118783764.wbieme0043

Villarino, R. T. (2025). Artificial Intelligence (AI) integration in Rural Philippine Higher Education. *IJERI: International Journal of Educational Research and Innovation*, *23*. https://doi.org/10.46661/ijeri.10909

Williamson, B., Komljenovic, J., & Gulson, K. (2023). *World Yearbook of Education 2024. Digitalisation of Education in the Era of Algorithms, Automation and Artificial Intelligence* . Routledge. https://doi.org/10.4324/9781003359722

Zain Abdillah, H., Partino, & Madjid, Abd. (2023). Enhancing Student Well-being through AI Chat GPT in the Smart Education University Learning Environment:

A Preliminary Review of Research Literature. *E3S Web of Conferences*, *440*, 05005. https://doi.org/10.1051/e3sconf/202344005005

Zawacki-Richter, O. (2024). The Origins of the Term Distance Education and the Roots of Digital Teaching and Learning. *Canadian Journal of Learning and Technology*, *49*(3), 1–7. https://doi.org/10.21432/cjlt28662

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7

Zhang, C., Rice, R. E., & Wang, L. H. (2024). College students' literacy, ChatGPT activities, educational outcomes, and trust from a digital divide perspective. *New Media and Society*. https://doi.org/10.1177/14614448241301741

Zhang, C., Wang, L. H., & Rice, R. E. (2025). U.S. college students' acceptability and educational benefits of ChatGPT from a digital divide perspective. *Computers and Education: Artificial Intelligence*, *8*. https://doi.org/10.1016/j.caeai.2025.100385

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