

Review

Deconstructing the Normalization of Data Colonialism in Educational Technology

Lucas Kohnke ¹  and Dennis Foung ^{2,*} 

¹ Department of English Language Education, The Education University of Hong Kong, Tai Po, Hong Kong; lmakohnke@eduhk.hk

² School of Journalism, Writing, and Media, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

* Correspondence: dennis.foung@gmail.com

Abstract: As learning analytics and educational data mining have become the “new normal” in the field, scholars have observed the emergence of data colonialism. Generally, data colonialism can be understood as the process by which data were considered “free” to take and appropriate. Building on this theoretical understanding, this study aims to contextualize data colonialism in educational technology by identifying and reviewing learning analytics studies that adopted a predictive analytics approach. We examined 22 studies from major educational technology journals and noted how they (1) see data as a resource to appropriate, (2) establish new social relations, (3) show the concentration of wealth, and (4) promote ideologies. We found evidence of data colonialism in the field of educational technology. While these studies may promote “better” ideologies, it is concerning how they justify the authorities capitalizing on “free” data. After providing a contextualized view of data colonialism in educational technology, we propose several measures to decolonize data practices, adopting a postcolonialist approach. We see data colonialism not only as a privacy issue but also as a culture that must be challenged.

Keywords: data colonialism; learning analytics; educational technology



Citation: Kohnke, L.; Foung, D. Deconstructing the Normalization of Data Colonialism in Educational Technology. *Educ. Sci.* **2024**, *14*, 57. <https://doi.org/10.3390/educsci14010057>

Academic Editors: Marguerite Koole, Matt Smith, John Traxler, Taskeen Adam and Shri Footring

Received: 25 October 2023
Revised: 11 December 2023
Accepted: 19 December 2023
Published: 3 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past three decades, the field of educational technology has expanded significantly and continually promised to transform education [1]. However, the “wow factor” associated with new technologies (e.g., radio, CD-ROMs, interactive whiteboards, virtual/augmented reality) often overshadows the actual needs of learners and leads to their uncritical acceptance or “normalization” [2]. Some digital technologies have even become as ubiquitous as traditional tools such as pens and paper [2]. Scholars have also begun to consider how the “normalization” of technology impacts both students and teachers.

Artificial intelligence (AI), which has existed for many years, has recently experienced a resurgence in popularity and interest due to the emergence of generative AI tools, such as ChatGPT. It has the potential to reshape teaching and learning once again [3,4] by optimizing face-to-face, blended, and online learning [5,6]. AI can retrieve large amounts of data from various sources, identify patterns, and cluster/predict these patterns; this constitutes its “intelligence”. Furthermore, software engineers deploy these patterns to perform human-like actions, which makes it “artificial”. AI-powered tools can assist educators in identifying and utilizing effective pedagogies based on learning data, generating teaching materials and assessments, and issuing grades and feedback automatically [7].

In education, the term “learning analytics” (LA) is typically used to describe the use of data to inform teaching. LA can be defined as the “measurement, collection, analysis, and reporting of data about learners and their contexts, for the purpose of understanding and optimizing learning and the environment in which it occurs” [8]. Using data to produce actionable insights has become a key goal of utilizing AI in education.

Nevertheless, the integration of AI and LA programs in education raises significant concerns related to the use of educational data. It prompts questions about the “normalization” of technology in education and its impact on culture and values [9]. Previous studies have explored privacy concerns related to learning data, specifically considering students’ perspectives. Ifenthlær and Schumacher [10], for example, found that students are not willing to share their personal information or the records of their behaviour online. Other studies have investigated how to respect privacy while deploying educational technologies and LA [11]. While this research offers important insights, studies on educational technology have yet to catch up with general AI research and the theory of “data colonialism”, which highlights the problematic nature of the massive retrieval and capturing of data.

The concept of “data colonialism” was introduced by Nick Couldry and Ulises Mejias, two scholars in Communications and Media Studies, and expounded in various publications, such as [12,13]. Working with other scholars (e.g., [14]), they identified similarities between colonialism and the data extraction practices of recent years. Under colonialism, natural resources were considered “free” to take and appropriate, which was supposed to bring about a new social order and a better world. Similarly, “data colonialism” treats user data as a natural resource, justifying the process by introducing new social relations and ideologies.

In a macro sense, Couldry and Mejias [12] draw on examples from major technology companies, such as Facebook and Amazon, which retrieve and privatize transaction data to connect user behaviours with personal attributes (i.e., social relations) and promote a more “personalized” purchasing experience (i.e., “a better world”). This leads to significant financial gains for these corporations and can convince customers to offer up more of their data (i.e., “free” resources). Even though critics such as Mumford [15] see this as a matter of data ownership, which could be addressed through regulations, the concept of data colonialism explains how companies are using seemingly “free” data for economic benefits.

Unfortunately, educational entities are not immune to such practices. Zembylas [16] has attempted to further contextualize how AI and LA can introduce data colonialism into higher education. In the context of educational technology, learning data are commonly used to generate value (though not always profit) for institutions and promote personalized learning experiences. Moreover, users are not always aware that their data have been appropriated. Thus, the concept of data colonialism in this context deserves further exploration.

Even though data colonialism is an important notion that has raised concerns in the academic community, to date, only a few conceptual discussions (such as [12,14,17,18]) have emerged. Little research has put data colonialism in context or examined how data are being appropriated. This study aims to provide a preliminary review of the realization of data colonialism in the field of educational technology. It is not intended to provide a comprehensive and in-depth synthesis but a general overview.

In addition, this preliminary review will not immediately provide solutions or identify how to “decolonize” educational technology. However, in response to Zembylas [16], it represents the first step of this process. By providing context and evidence, it can initiate a conversation about adopting “decolonized” practices in educational technology.

2. Methodology

This study aimed to review the existing body of literature on data colonialism by reviewing articles from impactful journals on educational technology. After choosing four journals, we conducted an initial search to select articles that related directly to our discussion. We then examined how the following four key features of data colonialism are being realized: (1) appropriation of resources; (2) establishing social relations; (3) concentration of wealth; and (4) promotion of ideologies. This allowed us to provide an overview of the topic.

2.1. Search Strategy

We chose “predictive analytics” as our search keyword because there are many studies on LA using educational technology, and this is one of the core research areas [19]. Using LA to predict student success—with the help of educational technology products and other solutions offered by vendors—is commonplace in higher education [20,21]. This keyword allowed us to identify many articles about LA.

We narrowed our focus to impactful journals by identifying the top five journals about educational technology on Google Scholar and Scopus, as well as all educational technology journals indexed in the Web of Science Social Science Citation Index (SSCI). When we examined these three lists, four journals appeared twice: *Education Technology & Society*, *International Journal of Educational Technology in Higher Education*, *British Journal of Educational Technology*, *Educational Technology Research and Development*, and *Australasian Journal of Educational Technology*. Therefore, we focused on articles published in these journals. Figure 1 presents details regarding how studies were included and/or excluded throughout this process.

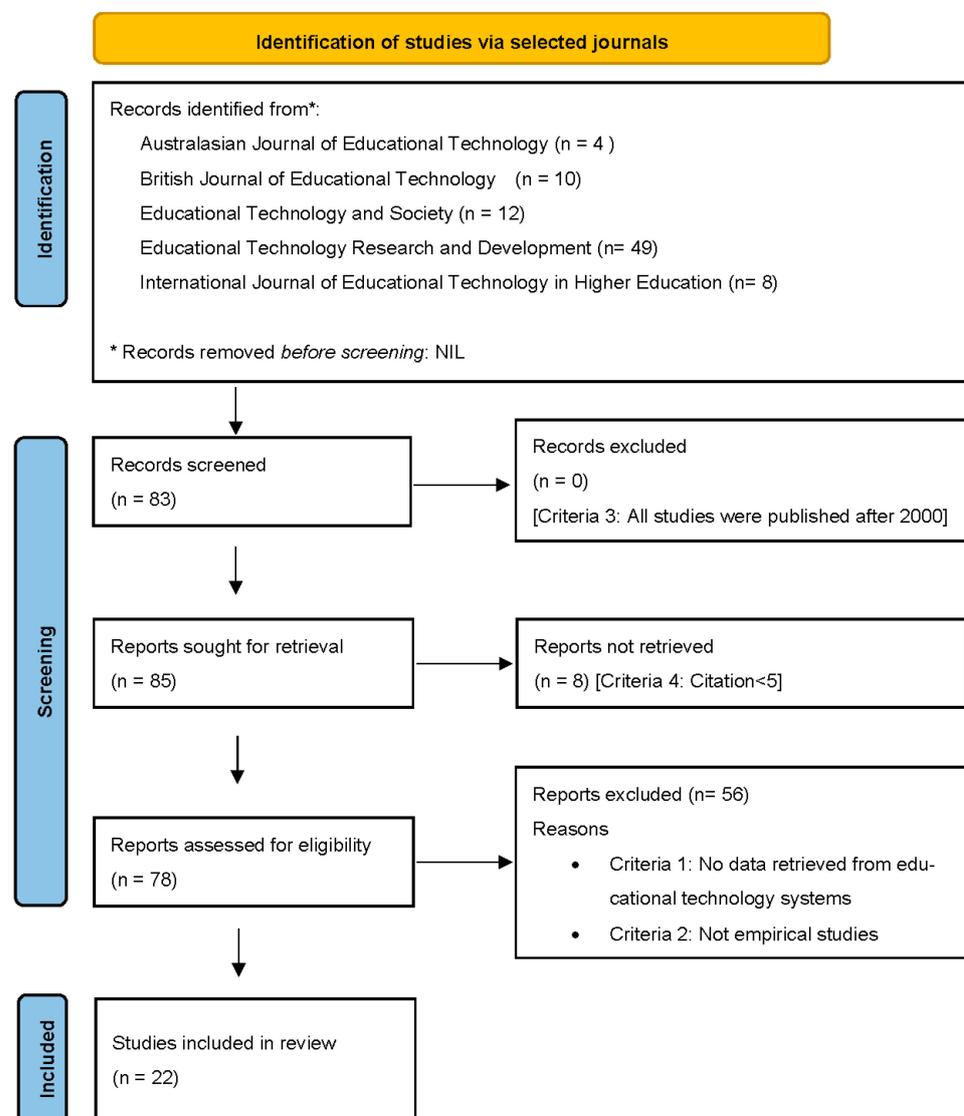


Figure 1. PRISMA 2020 flow diagram on article identification (adapted from Page et al. [22]).

2.2. Article Identification

After conducting our initial search, we identified a total of 83 studies with no duplicates. As we had targeted specific journals, all of the papers were peer-reviewed and written in English. We then applied the following inclusion criteria: (1) empirical studies; (2) data retrieved from an educational technology system (i.e., LA); (3) published after the year 2000; and (4) more than five citations.

Some of these criteria deserve brief explanations. The second criterion allowed us to exclude studies with traditional data collection strategies, such as questionnaires or interviews (which participants consent to complete). Because such participants provided their data willingly, these studies did not fit our aim. Using the fourth criterion, we ensured that we only included studies that have already received some attention in the field. While we believe that all of the studies in these journals are high quality, studies with at least five citations have gained recognition from the scholarly community, making them our priority.

2.3. Data Analysis

To understand how data colonialism is being realized in educational technology research, we examined four of its key features (see [12] for a detailed account of the concept). We developed key questions to correspond to each feature, as presented in Table 1.

Table 1. Guiding questions for data extraction.

Feature of Data Colonialism	Guiding Question
Appropriation of Resources	What data are being retrieved?
Social Relations	Other than the data being retrieved, what other information about users is involved?
Concentration of Wealth	Who has the privilege to approve the use of data? Are users aware that their data are being retrieved?
Promotion of Ideologies	What “better” outcome is being presented as the result of using the data?

3. Results

3.1. Overview of Studies

The final dataset included 22 studies published between 2013 and 2023. The number of citations in the studies (as of 1 October 2023) ranged from 6 [23] to 146 [24]. Among these studies, two were from the *Australasian Journal of Educational Technology*, four were from *Educational Technology Research and Development*, and eleven were from *Educational Technology Research and Development*. No studies from *Education Technology & Society* were included upon considering the inclusion and exclusion criteria. A general summary of the sources identified can be found in Appendices A and B. As the studies came from educational technology journals, most focused on learning behaviours or the effectiveness of particular platforms. They included studies on learning argumentation [25,26], facilitating academic advising sessions [27], and coding for kids [28]. Their samples ranged from less than 50 [25,29–32] to more than 100,000 students [24,28]. Many were based on introducing a new educational technology program in either an undergraduate [25,33–38] or postgraduate course [28,38]. Other contexts included elementary/high school [26,38,39], professional development for teachers [23,40] or university academic advisors [27], and online programs [28,30,31,35]. Seven studies were from the United States [25,26,30,31,35,39], and four were from Australia [34,38,41,42]. Other studies were performed in Asia [37,43], the United Kingdom [23,35], and Ecuador [27]. One was conducted online and did not specify the location or demographics [28]. Five [29,32,33,36,38] did not explicitly disclose the location despite being empirical studies.

3.2. Features of Data Colonialism

The following section describes the features of data colonialism identified in the studies based on the guiding questions presented in the previous section. After each feature is introduced, it is discussed with reference to the literature.

3.2.1. Appropriation of Resources

Among the studies reviewed, most retrieved behavioural data that had been generated by users of an educational technology system, including game logs [40], page views [24,39], and usage of an e-book tool [33] or learning management system [31,34]. Some studies were interested in user interaction data, such as forum posts [30,31] or chatroom chat logs [43]. Others were interested in spatial data and adopted tracking devices to capture and exploit the movements of learners [38,44]. A few retrieved assignments [25,26,29]. Importantly, all of these data were generated for other purposes (e.g., using a learning tool), not specifically for the research. They were then repurposed to promote the ideologies of the researchers. While many researchers captured log-based data, they also captured other data for linking purposes (e.g., questionnaire data or student outcome data). These data are described in the following section.

3.2.2. Social Relations

In these studies, log-based data were most often linked with questionnaire data. Researchers retrieved individuals' log-based data (as described in the previous section) and connected them to their answers on a questionnaire. The data included students' and teachers' strategies [42], affective outcomes [34], and experiences [39]. Log-based data were also linked with learning achievements, such as final grades [20,32–34,39,42], language test results [43], tests of concepts [25], and teachers' assessments [40,44]. Finally, log-based data were linked with teacher and student demographic data [23,33,36].

3.2.3. Concentration of Wealth

When data are considered a form of wealth, it is necessary to consider who has the power to distribute this wealth. In all of the studies, data produced by users for other purposes were appropriated for LA. While teachers (who may double as researchers) and IT departments can always access such data, we investigated the procedures by which the researchers obtained the authority to access this "wealth". Several studies did not disclose how they obtained approval to retrieve the data [26,34,37]. Unsurprisingly, most stated that an institutional research board or ethical clearance committee was able to approve this access through a data request [23,30–33,35,39–42]. Some studies, however, indicated that approval was "not required" [40], with one claiming that approval was "not applicable" because of "the nature of a study conducted on already available/existing data" [29]. This reflects the notion that wealth is "just there" to be capitalized on by others. It is encouraging to see that a few studies gave the power back to users and obtained their informed consent to use the data they produced [27,44].

3.2.4. Promotion of Ideologies

The ideologies promoted in the reviewed studies were consistent. Most were concerned with engagement [30,31,41,43], outcomes [23,29,35,39,43], or experiences [27,39,41]. Some were more specific, considering how to adopt educational technologies effectively [33]. While these ideologies are noble, other researchers with access to the same data may not share these aims. While these ideologies may also exist in other research disciplines, their use as an excuse to exploit data matches the notion of data colonialism.

4. Discussion

4.1. The Presence of Data Colonialism and Related Concerns

The results of this study suggest that data colonialism exists in educational technology research. In general, data were produced using private tools [28,35], higher education

learning management systems [31,34], and location tracking tools [38,44]. They were then captured and repurposed by researchers, including teachers [25,26] and members of the general public [28]. While researchers may have had admirable intentions, such as improving engagement [30,31,41,43], outcomes [23,29,35,39,43], or experiences [27,39,41], users were not always given a chance to agree to the use of their data. In practice, some users were only informed that their data was being used [43], and many were not even aware of this because approval was granted by ethics committees [30,31,33,35,39–42]. This practice echoes the idea that data simply exists, and anyone can take advantage of this [12]. Furthermore, under capitalism, no one can control whether such data will be exploited by others with different, less noble intentions. Some data from these studies are publicly available [28], so future researchers or private contractors will be able to capitalize on it without being bound by any constraints.

Our results highlight three major concerns related to data colonialism. First, data colonialism can further marginalize particular communities of learners. When researchers use existing data to establish new relationships with demographic variables [23,33,36] or final grades [34,39,40,42], they also establish relationships between students' demographics (e.g., race and gender) and behaviours. We found that these patterns may change more often in education than in other fields. For example, Williams et al. [36] examined students' use of a lecture-capturing podcast and concluded that Asian students and women were the heaviest users at this particular US university. Asian women were then chosen for further discussion in the study; the results of students from other races were not further discussed. In their study, the authors specifically focused on Asian women and found that heavy usage did not correlate with exam performance. This conclusion was drawn without making a comparison to other groups. While the study makes the argument that its findings are meant to relate to other literature, this can be considered as the first step of marginalizing Asian women. If these marginalized communities are targeted, their learning experience may be affected in the future. It is possible that some teachers would neglect heavy usage as an indicator of diligence based on the results of this study, making students feel that their time was wasted.

Second, the power dynamics between teachers, educational technology researchers, and learners make educational data especially vulnerable to data colonialism. For example, in relation to marketing analytics or social media analytics, users can choose not to use certain platforms to prevent their data from being colonized (as suggested by [13]). However, in educational institutions, it is hard for users to refuse. In practice, students generate data through courses they have to take for credit [25,32–37,42,43]. This may involve an educational technology tool they are required to use to pass the course or complete a mandatory assignment. The data students generate can then be retrieved for research purposes, a practice that can be seen as a form of colonial aggression.

In this context, teachers and/or educational technology researchers can also leverage their roles to require students to generate data/wealth, which can then be retrieved and capitalized upon. Significantly, this process also contributes to the advancement of researchers who benefit from the extraction of this "data-wealth". After obtaining approval from educational institutions [32,33,35,42], the data can be repurposed and exploited, often without giving students a chance to refuse or informing them that their data have been retrieved. This scenario can occur only because educational administrators or teachers hold power over their students, creating an unbalanced relationship that closely resembles colonialism. Therefore, these users are especially vulnerable to data colonialism and the exploitation of their data to benefit others.

Third, it is also concerning to consider the data retrieval and approval process. We have identified six levels of data sovereignty, from studies with no information on how they obtained approval for data retrieval to those giving users a choice of whether to participate. At the lowest level, some studies do not even disclose how the data retrieval was approved [26,34,37]. Other studies claimed prior approval was not required or necessary [29,35] but still disclosed this practice. At level three, one study used a secondary

dataset available online [28]. Many studies followed a conventional approach and gained access to data after ethics clearance from institutions [30,31,33,35,39–41]. At this level, students may still not know that their data are being retrieved or used for research purposes. At level five, one study informed students that they were using their data [43], which we consider a better practice. At the highest level, many studies asked students for explicit consent [23,25,27,32,36,38,42,44]. This practice provides students a chance to agree or disagree with the use of their data. These six levels of data retrieval practices provide a contextual overview of how data colonialism takes place in the field of educational technology. In subsequent sections, we offer recommendations to decolonialize such data retrieval practices.

4.2. Limitations

While we position the current study as an exploratory overview of the current literature, several limitations deserve readers' attention. First, it is ironic for the current review to choose only studies from the most impactful journals to examine colonialism. This means studies that embraced the "English language and Euro-Western worldviews" [45], which is made apparent by the notions of "better" and "more effective" in the reviewed studies. Unfortunately, this is a common issue in systematic reviews. (See de Almeida and de Goulart [46] for more discussion.) We believe that this review is only a starting point for understanding the so-called "mainstream" literature; more can be done afterwards.

Second, only one search term (i.e., "predictive analytics") was used to represent the field of learning analytics. The original intention was to gather any studies on data-driven analytics (see inclusion criteria), as predictive analytics is an important stream of research within the field. We eventually included all data-driven studies, which may have excluded other important learning analytics studies (e.g., those that profiled students through clustering). In other words, the studies identified are not yet representative of all learning analytics studies.

4.3. Implications and Recommendations

After finding evidence of data colonialism in education technology research, it is difficult to decide what to do next. User data generated by educational technology are available, their use is endorsed by institutions, and researchers take advantage of them to promote their ideologies. While we can offer some suggestions to empower the "colonized" users of educational technology, we are reluctant to argue that researchers must stop retrieving or mining data as a form of "decolonization". LA, AI, and educational data mining have established positions in the world of knowledge.

However, it may be possible to perceive data colonialism through a traditional post-colonialist lens. Postcolonialism generally refers to the study of formerly colonized cultures [46]; it often refers to hybridity, as suggested by Bhabha [47,48], and acknowledges the value of both the identities and knowledge that are produced through the process of colonization and those that pre-dated it. This notion of "hybridity" has started to emerge from the technological literature (e.g., [49]). Such an approach may help us move forward from arguing that data colonialism exists to embracing the postcolonialist world. In practice, we propose the following steps to decolonize data practices:

1. Respecting data sovereignty: Institutional ethics committees need to ensure that researchers have made a reasonable attempt to decolonialize their data practices by obtaining consent from users before using their data. While this is not always possible, especially with large institutional datasets, this review shows that it is sometimes possible to obtain student consent. In our review, we acknowledge that Yan et al. [38] and Broadbent and Fuller-Tyszkiewicz [42] did ask for consent from users despite retrieving their data directly from the university computer systems. This shows a significant effort to respect users' "right to be forgotten" [50].
2. Sensible data relations building: Institutional ethics committees should decolonialize their review of data retrieval requests and consider how researchers are building

relationships between variables. Only theoretically or empirically meaningful relationships should be examined. In our review, we were pleased that behavioural data were seldom linked to demographic data, as this is one of the students' major concerns (see [10]). If there are too many linkages or data points, ethics committees should be cautious about how this could affect the personal lives of users, especially those from marginalized communities.

3. **Avoiding manipulation of user behaviours:** We do not dispute the ideologies promoted by the reviewed studies, such as promoting engagement [30,31,41,43] or analyzing the effectiveness of programs [30,31]. To embrace a postcolonialist perspective, however, knowledge derived from data analytics alone should be deployed with caution. First, educational technology practitioners should further their understanding of user behaviour based on self-reported measures [30,31,33,34,39,42] or qualitative approaches [27]. Second, measures that aim to promote engagement or improve outcomes should not manipulate users' behaviour.
4. **Decolonizing the ethical clearance process:** While ethics clearance committees do not usually include students due to their technical and academic nature, institutions should consider engaging students, staff members, and other users in approving data retrieval requests. We believe that the best practice is to ask for consent directly. If that is impossible or inappropriate due to the ecology of ethics approval at an institution, one appropriate first step towards decolonization would be to include student members in the data retrieval committee, which approves and rejects requests from researchers. Having all data users represented can provide a sense of "sensible relationship building" and "avoiding manipulation of behaviours" described above.
5. **Decolonizing system design:** While we do not have the technical knowledge necessary, we suggest decolonizing educational technology systems from the top down (i.e., the system design level). Modern university systems are linked together, and user attributes are shared among databases. For example, students' numbers and preferred names are entered into the registrar's system and shared with the learning management system. In recent decades, educational institutions have adopted the inclusive practice of allowing users to enter their preferred pronouns on various systems (see [51] for a detailed discussion). We argue that institutions could also permit users to choose whether their data are shared across systems. With this attribute, IT personnel could retrieve data after filtering out those who have exercised their "right to be forgotten". Instead of retrieving all user data and deidentifying it manually, omitting data from certain users may be a more decolonized practice.
6. **Informing students about data use:** As part of the data consent process, students should be informed at the point of registration that the data they generate by interacting with the institution's systems may be utilized for various purposes. This can include not only the improvement of courses and programmes but also research purposes. This transparency could empower students to make informed decisions about their data and contribute to the decolonization of data practices.

5. Conclusions

Colonization has never been alien to the educational community, and this study shows that it is manifesting in the use of data for research, as well. This review study examined 22 articles using a predictive analytics approach and educational technology data. We found that data colonialism is common in the field of educational technology. With vulnerable data users and administrators who are in an "ivory tower," educational technology produces a broad range of data that is "just there" to be exploited. Promising better learning outcomes, researchers retrieve, repurpose, and link data. While some users were fortunate enough to have control over their data, others' data were used based on the approval of institutional ethics committees.

We are concerned that this sort of data colonialism could lead to the further marginalization of some learners. However, we are not advocating for researchers to stop using data

completely in order to achieve the “decolonization” of educational technology. Instead, we have proposed a range of measures to decolonize data practices so users can regain data sovereignty and limit their chances of being manipulated by algorithms. These practices may not fully decolonize educational technology, but they can at least raise awareness of data colonization.

Author Contributions: Conceptualization, L.K. and D.F.; methodology, D.F.; formal analysis, D.F. and L.K.; writing—original draft preparation & review and editing, D.F. and L.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. General summary of 22 studies being reviewed.

Citation Entry (#)	Article Title	Year	Authors	Links (All Accessed on 26 September 2023)
[22]	Analysis of patterns in time for evaluating effectiveness of first principles of instruction	2022	Frick et al.	https://link.springer.com/article/10.1007/s11423-021-10077-6
[23]	A large-scale implementation of predictive learning analytics in higher education: the teachers’ role and perspective	2019	Herodotou et al.	https://link.springer.com/article/10.1007/s11423-019-09685-0
[24]	The effects of successful versus failure-based cases on argumentation while solving decision-making problems	2013	Tawfik and Jonassen	https://link.springer.com/article/10.1007/s11423-013-9294-5
[25]	Identifying patterns in students’ scientific argumentation: content analysis through text mining using Latent Dirichlet Allocation	2020	Xing et al.	https://link.springer.com/article/10.1007/s11423-020-09761-w
[26]	Adoption and impact of a learning analytics dashboard supporting the advisor—Student dialogue in a higher education institute in Latin America	2020	De Laet et al.	https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12962
[27]	Understanding the relationship between computational thinking and computational participation: a case study from Scratch online community	2021	Jiang et al.	https://link.springer.com/article/10.1007/s11423-021-10021-8
[28]	To design or to integrate? Instructional design versus technology integration in developing learning interventions	2020	Kale et al.	https://link.springer.com/article/10.1007/s11423-020-09771-8
[29]	Priming, enabling and assessment of curiosity	2019	Sher et al.	https://scholar.google.ca/scholar?hl=en&as_sdt=0,5&q=Priming,+enabling+and%C2%A0assessment+of%C2%A0curiosity&btnG=
[30]	Exploring indicators of engagement in online learning as applied to adolescent health prevention: a pilot study of REAL media	2020	Ray et al.	https://link.springer.com/article/10.1007/s11423-020-09813-1

Table A1. Cont.

Citation Entry (#)	Article Title	Year	Authors	Links (All Accessed on 26 September 2023)
[31]	Gamification during COVID-19: Promoting active learning and motivation in higher education	2021	Rincon-Flores and Santos-Guevara	https://ajet.org.au/index.php/AJET/article/view/7157
[32]	The adoption of mark-up tools in an interactive e-textbook reader	2016	Van Horne et al.	https://link.springer.com/article/10.1007/s11423-016-9425-x
[33]	Academic success is about self-efficacy rather than frequency of use of the learning management system	2016	Broadbent	https://ajet.org.au/index.php/AJET/article/view/2634
[34]	Empowering online teachers through predictive learning analytics	2019	Herodotou et al.	https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12853
[35]	Lecture capture podcasts: differential student use and performance in a large introductory course	2015	Williams et al.	https://link.springer.com/article/10.1007/s11423-015-9406-5
[36]	Learning Analytics at Low Cost: At-risk Student Prediction with Clicker Data and Systematic Proactive Interventions	2018	Choi et al.	https://www.jstor.org/stable/26388407
[37]	The role of indoor positioning analytics in assessment of simulation-based learning	2022	Yan et al.	https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.13262
[38]	Predict or describe? How learning analytics dashboard design influences motivation and statistics anxiety in an online statistics course	2021	Valle et al.	https://link.springer.com/article/10.1007/s11423-021-09998-z
[42]	Do social regulation strategies predict learning engagement and learning outcomes? A study of English language learners in wiki-supported literature circles activities	2021	Li et al.	https://link.springer.com/article/10.1007/s11423-020-09934-7
[39]	Does slow and steady win the race?: Clustering patterns of students' behaviors in an interactive online mathematics game	2022	Lee et al.	https://link.springer.com/article/10.1007/s11423-022-10138-4
[43]	Mapping from proximity traces to socio-spatial behaviours and student progression at the school	2022	Yan et al.	https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.13203
[41]	Profiles in self-regulated learning and their correlates for online and blended learning students	2018	Broadbent and Fuller-Tyszkiewicz	https://link.springer.com/article/10.1007/s11423-018-9595-9
[40]	Identifying engagement patterns with video annotation activities: A case study in professional development	2018	Mirriahi et al.	https://ajet.org.au/index.php/AJET/article/view/3207

Appendix B

Table A2. General summary of 22 studies being reviewed.

#	Authors	No. of Citations (as at 1 October 2023)	Context	Sample Size	Location	Data Retrieved from Educational Technology Systems	Other Data Collected/Retrieved (Excepted)
[22]	Frick et al.	6	University Teachers	59	UK	Login data of a dashboard	<ul style="list-style-type: none"> • Student final grades • Teacher demographic info (gender) • Student demographic info
[23]	Herodotou et al.	146	MOOC	172,417	US	Usage data on webpages (pageviews, clicks scrolling)	nil
[24]	Tawfik and Jonassen	85	Undergraduate	36	US	Arguments produced by users	Pretest and post-test of concepts
[25]	Xing et al.	26	Middle/High School	2472	US	Student produced arguments	Teacher assessment of students' learning
[26]	De Laet et al.	34	University Academic Advisors	172	Ecuador	Student study plan before and after intervention	Simulated advising sessions (qualitative data)
[27]	Jiang et al.	10	online learning tool	105,720	Online	Online learning journey (likes/loves) remixing projects	Computation scores assigned by another researcher
[28]	Kale et al.	17	Postgraduate	22	Not mentioned	Final projects completed for courses	nil
[29]	Sher et al.	11	Online program for youth club	38	US	Participant interactions	Questionnaire data on audience engagement
[30]	Ray et al.	13	Online Substance use prevention program	38	US	User interactions on the LMS	Questionnaire data on program usability
[31]	Rincon-Flores and Santos-Guevara	54	Undergraduate	40	Not mentioned	Student final grades and course achievement	Student grade
[32]	Van Horne et al.	71	Undergraduate	274	"Midwest"	Student Usage of mark-up tool (for a reading tool)	Questionnaire on reading behaviour
[33]	Broadbent	100	Undergraduate	310	Australia	Student LMS usage data	Questionnaire data on self-efficacy locus of control motivation
[34]	Herodotou et al.	79	Undergraduate	559	UK	Usage of dashboard system	Discipline of teachers/student performance
[35]	Williams et al.	46	Undergraduate	835	not mentioned	Login data from video viewing site	In-class clickers student demographic
[36]	Choi et al.	113	Undergraduate	1075	Hong Kong	In-class clickers data	Demographic information

Table A2. Cont.

#	Authors	No. of Citations (as at 1 October 2023)	Context	Sample Size	Location	Data Retrieved from Educational Technology Systems	Other Data Collected/Retrieved (Excepted)
[37]	Yan et al.	12	Undergraduate	3604	Australia	Position tracking in a simulated room	Teacher assessment of students' learning
[38]	Valle et al.	20	Postgraduate	179	US	Number of views	Questionnaire data on prior content knowledge, experience
[42]	Li et al.	20	English language course	95	China	QQ chatroom chat logs	Language test at the end of activities
[39]	Lee et al.	9	Middle school	227	US	Student game logs	<ul style="list-style-type: none"> • Type of math classes attending • Gender • Grade
[43]	Yan et al.	8	Elementary	98	Not mentioned	Position tracker/wearable device position data	Student progression
[41]	Broadbent and Fuller-Tyszkiewicz	122	Undergraduate	606	Australia	Final grade	<ul style="list-style-type: none"> • Questionnaire data on motivational and self-regulated learning strategies • Student demographic information
[40]	Mirriahi et al.	37	Teachers	163	Australia	Behavioural data on video annotation tool	nil

References

- Laurillard, D. Supporting Teachers in Optimizing Technologies for Open Learning. In *Global Challenges and Perspectives in Blended and Distance Learning*; IGI Global: Hershey, PA, USA, 2013; pp. 160–173, ISBN 978-1-4666-3978-2.
- Bax, S. CALL—Past, Present and Future. *System* **2003**, *31*, 13–28. [[CrossRef](#)]
- Kohnke, L.; Moorhouse, B.L.; Zou, D. ChatGPT for Language Teaching and Learning. *RELC J.* **2023**, *54*, 537–550. [[CrossRef](#)]
- Kohnke, L.; Moorhouse, B.L.; Zou, D. Exploring generative artificial intelligence preparedness among university language instructors. *Comput. Educ. Artif. Intell.* **2023**, *5*, 100156. [[CrossRef](#)]
- Chiu, T.K.F.; Xia, Q.; Zhou, X.; Chai, C.S.; Cheng, M. Systematic Literature Review on Opportunities, Challenges, and Future Research Recommendations of Artificial Intelligence in Education. *Comput. Educ. Artif. Intell.* **2023**, *4*, 100118. [[CrossRef](#)]
- Kexin, L.; Yi, Q.; Xiaoou, S.; Yan, L. Future Education Trend Learned From the COVID-19 Pandemic: Take «Artificial Intelligence» Online Course As an Example. In Proceedings of the 2020 International Conference on Artificial Intelligence and Education (ICAIE), Online, 26–28 June 2020; pp. 108–111.
- Chaudhry, M.A.; Kazim, E. Artificial Intelligence in Education (AIED): A High-Level Academic and Industry Note 2021. *AI Ethics* **2022**, *2*, 157–165. [[CrossRef](#)] [[PubMed](#)]
- Society for Learning Analytics Research What Is Learning Analytics? Available online: <https://www.solaresearch.org/about/what-is-learning-analytics/> (accessed on 22 October 2023).
- Mhlambi, S. Decolonizing AI. Available online: <https://www.youtube.com/watch?v=UqVwfuluU2k&t=30s> (accessed on 1 October 2023).
- Ifenthaler, D.; Schumacher, C. Student Perceptions of Privacy Principles for Learning Analytics. *Educ. Technol. Res. Dev.* **2016**, *64*, 923–938. [[CrossRef](#)]
- Scholes, V. The Ethics of Using Learning Analytics to Categorize Students on Risk. *Educ. Technol. Res. Dev.* **2016**, *64*, 939–955. [[CrossRef](#)]
- Couldry, N.; Mejias, U.A. The Decolonial Turn in Data and Technology Research: What Is at Stake and Where Is It Heading? *Inf. Commun. Soc.* **2023**, *26*, 786–802. [[CrossRef](#)]
- Couldry, N.; Mejias, U.A. Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject. *Telev. New Media* **2019**, *20*, 336–349. [[CrossRef](#)]
- Thompson, T.L.; Prinsloo, P. Returning the Data Gaze in Higher Education. *Learn. Media Technol.* **2023**, *48*, 153–165. [[CrossRef](#)]
- Mumford, D. Data Colonialism: Compelling and Useful, but Whither Epistemes? *Inf. Commun. Soc.* **2022**, *25*, 1511–1516. [[CrossRef](#)]
- Zembylas, M. A Decolonial Approach to AI in Higher Education Teaching and Learning: Strategies for Undoing the Ethics of Digital Neocolonialism. *Learn. Media Technol.* **2023**, *48*, 25–37. [[CrossRef](#)]
- Thatcher, J.; O’Sullivan, D.; Mahmoudi, D. Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data. *Env. Plan. D* **2016**, *34*, 990–1006. [[CrossRef](#)]
- Prinsloo, P. Data Frontiers and Frontiers of Power in (Higher) Education: A View of/from the Global South. *Teach. High. Educ.* **2020**, *25*, 366–383. [[CrossRef](#)]
- Sghir, N.; Adadi, A.; Lahmer, M. Recent Advances in Predictive Learning Analytics: A Decade Systematic Review (2012–2022). *Educ. Inf. Technol.* **2023**, *28*, 8299–8333. [[CrossRef](#)] [[PubMed](#)]
- Smithers, L. Predictive Analytics and the Creation of the Permanent Present. *Learn. Media Technol.* **2023**, *48*, 109–121. [[CrossRef](#)]
- Williamson, B. Policy Networks, Performance Metrics and Platform Markets: Charting the Expanding Data Infrastructure of Higher Education. *Br. J. Educ. Technol.* **2019**, *50*, 2794–2809. [[CrossRef](#)]
- Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)]
- Frick, T.W.; Myers, R.D.; Dagle, C. Analysis of Patterns in Time for Evaluating Effectiveness of First Principles of Instruction. *Educ. Technol. Res. Dev.* **2022**, *70*, 1–29. [[CrossRef](#)]
- Herodotou, C.; Rienties, B.; Boroowa, A.; Zdrahal, Z.; Hlosta, M. A Large-Scale Implementation of Predictive Learning Analytics in Higher Education: The Teachers’ Role and Perspective. *Educ. Technol. Res. Dev.* **2019**, *67*, 1273–1306. [[CrossRef](#)]
- Tawfik, A.; Jonassen, D. The Effects of Successful versus Failure-Based Cases on Argumentation While Solving Decision-Making Problems. *Educ. Technol. Res. Dev.* **2013**, *61*, 385–406. [[CrossRef](#)]
- Xing, W.; Lee, H.-S.; Shibani, A. Identifying Patterns in Students’ Scientific Argumentation: Content Analysis through Text Mining Using Latent Dirichlet Allocation. *Educ. Technol. Res. Dev.* **2020**, *68*, 2185–2214. [[CrossRef](#)]
- De Laet, T.; Millecamp, M.; Ortiz-Rojas, M.; Jimenez, A.; Maya, R.; Verbert, K. Adoption and Impact of a Learning Analytics Dashboard Supporting the Advisor—Student Dialogue in a Higher Education Institute in Latin America. *Br. J. Educ. Technol.* **2020**, *51*, 1002–1018. [[CrossRef](#)]
- Jiang, B.; Zhao, W.; Gu, X.; Yin, C. Understanding the Relationship between Computational Thinking and Computational Participation: A Case Study from Scratch Online Community. *Educ. Technol. Res. Dev.* **2021**, *69*, 2399–2421. [[CrossRef](#)]
- Kale, U.; Roy, A.; Yuan, J. To Design or to Integrate? Instructional Design versus Technology Integration in Developing Learning Interventions. *Educ. Technol. Res. Dev.* **2020**, *68*, 2473–2504. [[CrossRef](#)]

30. Sher, K.B.-T.; Levi-Keren, M.; Gordon, G. Priming, Enabling and Assessment of Curiosity. *Educ. Technol. Res. Dev.* **2019**, *67*, 931–952. [[CrossRef](#)]
31. Ray, A.E.; Greene, K.; Pristavec, T.; Hecht, M.L.; Miller-Day, M.; Banerjee, S.C. Exploring Indicators of Engagement in Online Learning as Applied to Adolescent Health Prevention: A Pilot Study of REAL Media. *Educ. Technol. Res. Dev.* **2020**, *68*, 3143–3163. [[CrossRef](#)]
32. Rincon-Flores, E.G.; Santos-Guevara, B.N. Gamification during COVID-19: Promoting Active Learning and Motivation in Higher Education. *Australas. J. Educ. Technol.* **2021**, *37*, 43–60. [[CrossRef](#)]
33. Van Horne, S.; Russell, J.; Schuh, K.L. The Adoption of Mark-up Tools in an Interactive e-Textbook Reader. *Educ. Technol. Res. Dev.* **2016**, *64*, 407–433. [[CrossRef](#)]
34. Broadbent, J. Academic Success Is about Self-Efficacy Rather than Frequency of Use of the Learning Management System. *Australas. J. Educ. Technol.* **2016**, *32*, 2634. [[CrossRef](#)]
35. Herodotou, C.; Hlosta, M.; Borooowa, A.; Rienties, B.; Zdrahal, Z.; Mangafa, C. Empowering Online Teachers through Predictive Learning Analytics. *Br. J. Educ. Technol.* **2019**, *50*, 3064–3079. [[CrossRef](#)]
36. Williams, A.E.; Aguilar-Roca, N.M.; O'Dowd, D.K. Lecture Capture Podcasts: Differential Student Use and Performance in a Large Introductory Course. *Educ. Technol. Res. Dev.* **2016**, *64*, 1–12. [[CrossRef](#)]
37. Choi, S.P.M.; Lam, S.S.; Li, K.C.; Wong, B.T.M. Learning Analytics at Low Cost: At-Risk Student Prediction with Clicker Data and Systematic Proactive Interventions. *J. Educ. Technol. Soc.* **2018**, *21*, 273–290.
38. Yan, L.; Martinez-Maldonado, R.; Zhao, L.; Dix, S.; Jaggard, H.; Wotherspoon, R.; Li, X.; Gašević, D. The Role of Indoor Positioning Analytics in Assessment of Simulation-Based Learning. *Br. J. Educ. Technol.* **2023**, *54*, 267–292. [[CrossRef](#)]
39. Valle, N.; Antonenko, P.; Valle, D.; Sommer, M.; Huggins-Manley, A.C.; Dawson, K.; Kim, D.; Baiser, B. Predict or Describe? How Learning Analytics Dashboard Design Influences Motivation and Statistics Anxiety in an Online Statistics Course. *Educ. Technol. Res. Dev.* **2021**, *69*, 1405–1431. [[CrossRef](#)] [[PubMed](#)]
40. Lee, J.-E.; Chan, J.Y.-C.; Botelho, A.; Ottmar, E. Does Slow and Steady Win the Race?: Clustering Patterns of Students' Behaviors in an Interactive Online Mathematics Game. *Educ. Technol. Res. Dev.* **2022**, *70*, 1575–1599. [[CrossRef](#)]
41. Mirriahi, N.; Jovanovic, J.; Dawson, S.; Gašević, D.; Pardo, A. Identifying Engagement Patterns with Video Annotation Activities: A Case Study in Professional Development. *Australas. J. Educ. Technol.* **2018**, *34*, 3207. [[CrossRef](#)]
42. Broadbent, J.; Fuller-Tyszkiewicz, M. Profiles in Self-Regulated Learning and Their Correlates for Online and Blended Learning Students. *Educ. Technol. Res. Dev.* **2018**, *66*, 1435–1455. [[CrossRef](#)]
43. Li, Y.; Chen, K.; Su, Y.; Yue, X. Do Social Regulation Strategies Predict Learning Engagement and Learning Outcomes? A Study of English Language Learners in Wiki-Supported Literature Circles Activities. *Educ. Technol. Res. Dev.* **2021**, *69*, 917–943. [[CrossRef](#)]
44. Yan, L.; Martinez-Maldonado, R.; Gallo Cordoba, B.; Deppeler, J.; Corrigan, D.; Gašević, D. Mapping from Proximity Traces to Socio-Spatial Behaviours and Student Progression at the School. *Br. J. Educ. Technol.* **2022**, *53*, 1645–1664. [[CrossRef](#)]
45. Chambers, L.A.; Jackson, R.; Worthington, C.; Wilson, C.L.; Tharao, W.; Greenspan, N.R.; Masching, R.; Pierre-Pierre, V.; Mbulaheni, T.; Amirault, M.; et al. Decolonizing Scoping Review Methodologies for Literature With, for, and by Indigenous Peoples and the African Diaspora: Dialoguing With the Tensions. *Qual. Health Res.* **2018**, *28*, 175–188. [[CrossRef](#)] [[PubMed](#)]
46. De Almeida, C.P.B.; De Goulart, B.N.G. How to Avoid Bias in Systematic Reviews of Observational Studies. *Rev. CEFAC* **2017**, *19*, 551–555. [[CrossRef](#)]
47. Wang, Y. The Cultural Factors in Postcolonial Theories and Applications. *JLTR* **2018**, *9*, 650. [[CrossRef](#)]
48. Bhabha, H. *The Location of Culture*; Routledge: London, UK, 1994.
49. Peralta, L.M.M. Resisting Techno-Orientalism and Mimicry Stereotypes in and Through Data Science Education. *TechTrends* **2023**, *67*, 426–434. [[CrossRef](#)]
50. Drachsler, H.; Greller, W. Privacy and Analytics—It's a DELICATE Issue. A Checklist for Trusted Learning Analytics. In Proceedings of the 6th Learning Analytics and Knowledge Conference 2016, Edinburgh, UK, 25–29 April 2016.
51. Chan, B.; Stewart, J.J. Listening to Nonbinary Chemistry Students: Nonacademic Roadblocks to Success. *J. Chem. Educ.* **2022**, *99*, 409–416. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.