sustainability

MDPI

Systematic Review

A Systematic Review of Postgraduate Programmes Concerning Ethical Imperatives of Data Privacy in Sustainable Educational Data Analytics

Mthokozisi Masumbika Ncube and Patrick Ngulube *

Department of Interdisciplinary Research and Postgraduate Studies, University of South Africa, Pretoria 0003, South Africa; ncubem.masumbika@gmail.com
* Correspondence: ngulup@unisa.ac.za

Abstract: This systematic review, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, investigated the intersection of data privacy, postgraduate educational data analytics (EDA), and sustainability. Existing literature focuses on general privacy concerns in EDA, neglecting the specific data collected and related risks in postgraduate programmes. This review addresses this gap by identifying data types used by higher education institutions in postgraduate initiatives and evaluating the adequacy of current ethical frameworks, particularly for sustainability goals. Recognising the lack of established best practices for balancing data utility and privacy, the review analyses privacy-preserving techniques. Through identifying key data types collected in postgraduate initiatives, evaluating existing ethical frameworks, and exploring privacy-preserving educational data analytics techniques, this study provided practical guidance for higher education institutions to navigate the challenges of balancing data utility and student privacy. The results suggest that higher education institutions can achieve sustainable data use by adopting a comprehensive approach that incorporates best practices, emerging technologies, and continuous monitoring to safeguard student privacy while leveraging the benefits of educational data analytics for achieving the Sustainable Development Goals.

Keywords: educational data analytics; data privacy; learning analytics; sustainability; systematic review

check for updates

Citation: Ncube, M.M.; Ngulube, P. A Systematic Review of Postgraduate Programmes Concerning Ethical Imperatives of Data Privacy in Sustainable Educational Data Analytics. *Sustainability* **2024**, *16*, 6377. https://doi.org/10.3390/su16156377

Academic Editor: Gazi Mahabubul Alam

Received: 14 May 2024
Revised: 20 July 2024
Accepted: 23 July 2024
Published: 25 July 2024

1. Introduction

The global community faces a significant challenge of achieving the ambitious goals outlined in the United Nations' Sustainable Development Goals (SDGs) by 2030. These multifaceted goals, encompassing issues like poverty eradication and climate change mitigation, necessitate a comprehensive approach [1]. The SDGs emphasise the importance of quality education (SDG 4) and responsible consumption and production patterns (SDG 12) [2]. Higher education institutions (HEIs) play a vital role in achieving these goals, increasingly utilising educational data analytics (EDA) to monitor progress and optimise resource allocation [3]. EDA offers immense potential to enhance educational experiences and contribute to a more sustainable future. For instance, HEIs can leverage EDA to analyse student behaviour patterns related to energy consumption in campus buildings and identify areas for improvement [4]. As such, EDA serves as a crucial tool for informing decision making, measuring progress, and pinpointing areas needing improvement [1]. However, harnessing student data for sustainability initiatives presents ethical concerns regarding data privacy, particularly within the context of postgraduate programmes. Lane [5] highlights the potential for the collection and analysis of educational data to reveal sensitive details about students' learning styles, academic progress, and areas of weakness. Within the realm of sustainable education, where students often grapple with complex social and environmental issues, additional privacy considerations arise. For instance, student data could inadvertently capture personal opinions on controversial topics like environmental policies [6]. This could potentially lead to a proliferation of self-censorship or a reluctance to fully engage in learning activities [6].

The deluge of big data has propelled data privacy to the forefront of ethical concerns. The collection, storage, and analysis of student data for educational data analytics (EDA) raises concerns about potential misuse and unauthorised disclosure of personal information [7]. Striking a balance between leveraging the power of EDA for sustainability initiatives and safeguarding student privacy is paramount to fostering trust within the educational ecosystem [8]. This is particularly crucial in postgraduate programmes, where students often engage in independent research on sensitive topics. Such research may generate data considered more private than undergraduate studies due to the nature of the inquiry [8]. Additionally, postgraduate students possess greater autonomy over their research direction, highlighting the critical importance of informed consent regarding data utilisation. The legal landscape surrounding student data privacy is a complex and evolving domain. Regulations like the General Data Protection Regulation (GDPR) in the European Union (EU) impose strict limitations on the collection, storage, and use of personal data [9]. Understanding and complying with these frameworks is essential for HEIs to develop responsible and compliant educational data analytics practices. However, the GDPR is not the sole regulation to consider. Broader principles for data protection are established by international regulations like the Council of Europe's Convention for the Protection of Individuals concerning Automatic Processing of Personal Data and its Additional Protocol (Convention 108) [10]. The landscape of African data privacy regulations is still under development, but several countries are enacting comprehensive data protection laws [8]. For HEIs operating in Africa or collaborating with postgraduate students from African countries, understanding these emerging regulations is crucial.

This systematic review aims to address three critical research gaps at the intersection of data privacy considerations, postgraduate EDA, and sustainability. Existing literature on EDA often addresses broad privacy issues but fails to comprehensively explore the specific data points collected within postgraduate programmes and the associated risks for these students. This review identified the types of data commonly utilised in HEI initiatives focused on postgraduate students. Ethical frameworks for educational data use might not fully capture the nuances of postgraduate education, particularly when sustainability is a core objective. This review examined how current frameworks manage the specific needs and sensitivities of postgraduate data within a sustainability context. There is a lack of established best practices for HEIs to balance the benefits of using student data for sustainability initiatives with protecting postgraduate student privacy. To address this, this review systematically analysed existing research on privacy-preserving educational data analytics techniques. By identifying and evaluating these best practices, the study aims to provide HEIs with practical guidance on implementing responsible and privacy-conscious data use in postgraduate EDA projects that contribute to achieving the SDGs, and, as a systematic review, also aims to contribute valuable insights to the evolving field of educational data analytics in a sustainability context. By addressing these research gaps, the study promotes the responsible use of data for achieving sustainability goals while safeguarding the privacy rights of postgraduate students.

Aligned with this overarching goal, the review pursued the following research objectives:

- To categorise the types and dimensions of sustainable educational data typically collected in postgraduate programmes by HEIs.
- To explore the potential privacy risks that may arise from the utilisation of student data for sustainable educational analytics within postgraduate programmes.
- To examine current practices and emerging frameworks employed by HEIs to ensure the sustainable privacy of student data within postgraduate programme educational settings.

2. Materials and Methods

To ensure transparency and methodological rigour, this systematic review adopted the guidelines outlined in the Preferred Reporting Items for Systematic Reviews (PRISMA 2020, Supplementary Materials) statement [11].

2.1. Systematic Search Strategy and Study Selection

A comprehensive search strategy guided the identification of relevant studies for this systematic review. Three academic databases were utilised, including ERIC, Scopus, and Web of Science, alongside two additional databases specifically focused on sustainability research (GreenFILE and Sustainability Science Collection). The search encompassed peer-reviewed journal articles, conference papers, and book chapters published in English. The search terms were carefully chosen to reflect the core areas of sustainability in postgraduate education, educational data analytics, and student privacy. Examples include "postgraduate education", "masters", "doctoral", "sustainable education", "educational data analytics", "learning analytics", "student privacy", and "data ethics". Boolean operators (AND, OR) were strategically combined to ensure the search captured relevant literature while maintaining a focus on sustainable data privacy within postgraduate education. This multi-faceted approach aimed to achieve a balance between comprehensiveness and precision in identifying relevant studies for this review.

To ensure a targeted and relevant literature search focusing on data privacy concerns, this review adopted the Population, Intervention, Comparison, Outcome (PICO) framework to develop the search strategy and establish eligibility criteria [12]. Following the PICO elements, studies were selected based on the following criteria:

- Population (P): Focus: Postgraduate students (Master's or Doctoral programmes) in education. Screening: Studies explicitly mentioning postgraduate students received priority during title and abstract screening. Studies solely focussed on other educational levels or broader student populations were excluded.
- Intervention (I): Focus: Use of educational data analytics in postgraduate education. Screening: Priority was given to studies investigating interventions utilising educational data analytics for student support or learning improvement within postgraduate education. Studies solely focussed on the technical aspects of educational data analytics without an educational application were excluded.
- Comparison (C): Focus: Comparison group or baseline data on student privacy considerations. Screening: Studies that compared the effectiveness of educational data analytics interventions about student privacy were prioritised. This could involve comparisons to traditional methods or control groups without educational data analytics, or studies examining how existing interventions addressed privacy concerns. Studies with baseline data on student privacy practices before the intervention were also considered.
- Outcome (O): Focus: Student privacy considerations and potential risks. Screening: Studies that explored potential privacy risks associated with educational data analytics in postgraduate education were prioritised. This could encompass concerns about data collection methods, storage practices, student anonymity, and the potential for self-censorship due to data analysis. To ensure the retrieved studies directly addressed the research topic, the PICO framework was complemented by the following additional eligibility criteria:
  - To encompass the most recent advancements in educational data analytics and their concomitant privacy concerns, this study included English-language publications spanning the period 2011–2024. The selection of 2011 as the starting point is deliberate, coinciding with the exponential surge in big data collection and analytics capabilities. This period witnessed a demonstrably increased focus by educational institutions on collecting and analysing student data in novel and ever-more comprehensive ways [6].

– Studies that provided detailed information regarding educational data analytics interventions and how they addressed student privacy were prioritised. This includes specifics on data types collected, anonymisation techniques, and student consent procedures.
– Empirical research articles that demonstrated a critical perspective on student privacy within the educational data analytics context of postgraduate education were preferred. These studies employed sound research methodologies and contributed to the understanding of balancing data-driven educational practices with student privacy in this domain.

This refined approach ensured the selection of relevant literature that directly addresses the interplay between educational data analytics, student privacy, and postgraduate education.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Figure 1) depicts the search strategy and selection process for the included studies (adapted from [11]). The initial database search yielded a total of 522 studies. To efficiently manage this extensive corpus and ensure a rigorous selection process, a two-stage approach was adopted. This involved utilising Rayyan [13], a web-based platform designed for systematic reviews, alongside ASReview [14], an open-source software incorporating machine learning functionalities.

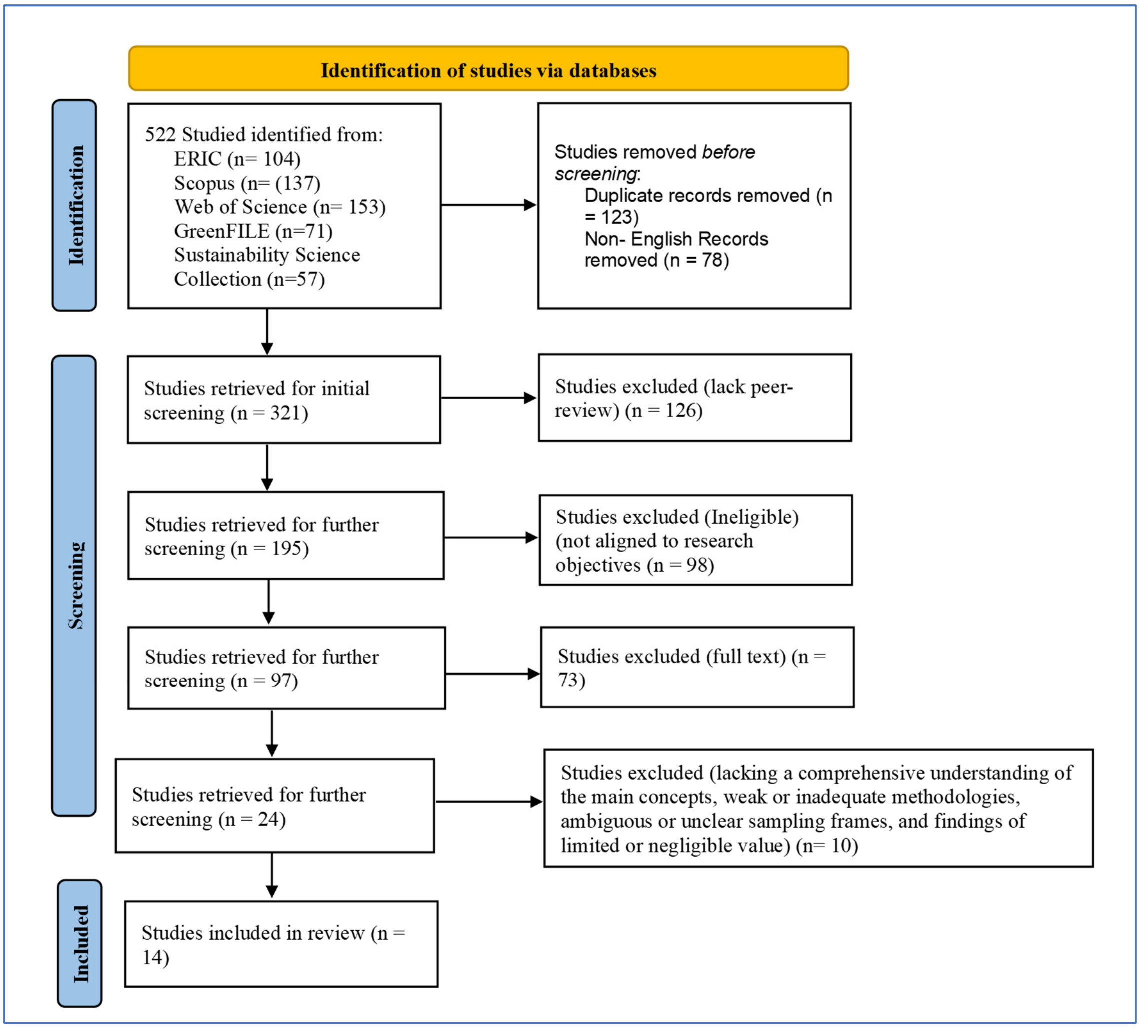

Figure 1. Mapping the Literature Search and Study Selection Process (adapted from PRISMA Statement [11]).

- Stage 1: De-duplication and Machine Learning Prioritisation
  - De-duplication: The initial stage commenced with leveraging Rayyan's [13] deduplication capabilities. This effectively removed 201 duplicate and non-English studies, resulting in a more manageable pool of 321 articles for further evaluation.
  - Machine Learning Prioritisation: Subsequently, ASReview [14], a tool powered by machine learning, analysed the remaining titles and abstracts. Based on the pre-defined inclusion and exclusion criteria established within the platform, ASReview prioritised articles for full-text review. This potentially reduced workload by focusing on the most relevant studies first.
- Stage 2: In-depth Review and Selection within Rayyan
  - Rigorous Full-Text Review: Following ASReview's [14] prioritisation, both researchers (reviewers) independently assessed the full text of the selected articles within the Rayyan [13] interface. This meticulous process involved applying the pre-defined eligibility criteria to identify studies that fully met the research objectives and aligned with the specific focus on data privacy concerns in postgraduate education.
  - Exclusion Based on Eligibility Criteria: Through this in-depth evaluation, 297 studies were excluded for not comprehensively satisfying the established criteria. This rigorous selection process resulted in the identification of twenty-four key studies directly addressing sustainable educational data analytics in postgraduate education with a focus on student privacy considerations.

2.2. Quality Assessment Using the CASP Checklist

The twenty-four studies were retained for a comprehensive quality assessment, focusing on the following key aspects:

- Clarity and Comprehensiveness of Privacy Concepts: The Critical Appraisal Skills Programme (CASP) [15] checklist guided the reviewers in evaluating the authors' understanding of relevant data privacy concepts, specifically within the context of postgraduate education. The reviewers assessed the introduction and literature review sections for evidence of a clear understanding of student privacy principles, potential privacy risks associated with educational data analytics in this domain, and existing research on mitigating these risks.
- Methodological Rigour in Examining Privacy Concerns: The CASP checklist [15] provided a structured framework for scrutinising the methodology sections. The reviewers examined factors like the appropriateness of research designs for the specific research questions related to data privacy, data collection methods employed (e.g., surveys addressing student concerns about data use), data analysis techniques that ensured student anonymity, and the researchers' attention to potential bias and its mitigation strategies. Studies employing well-justified and rigorous methodologies for investigating data privacy concerns were prioritised.
- Clear Sampling Frames and Representation: The CASP checklist emphasised the importance of clear sampling frames. The reviewers evaluated whether the authors clearly defined the target postgraduate student population, the sampling method used to select participants for studies on data privacy, and the justification for the sample size and its representativeness of the broader postgraduate student body. Studies with well-defined and representative sampling frames were considered more reliable for informing generalisable conclusions about data privacy practices.
- Valuable Findings on Data Privacy for Sustainable Education: The CASP checklist focuses on the results and discussion sections regarding data privacy. The reviewers assessed the clarity of the presented findings on student privacy concerns related to sustainable educational data analytics in postgraduate education, their alignment with the research questions and methodology, and the depth of discussion regarding their meaning and implications for developing responsible data governance practices.

Articles presenting valuable contributions to the knowledge base on data privacy in this specific context, along with practical recommendations for mitigating privacy risks, were deemed particularly noteworthy.

Through utilising the CASP checklist [15] as a guiding framework, a systematic and thorough evaluation of each article's strengths and weaknesses across these four key quality indicators was ensured. This approach led to the selection of a final collection of fourteen high-quality research articles that formed the foundation for the comprehensive review of data privacy considerations in sustainable educational data analytics for postgraduate education. Leveraging the PRISMA framework, Figure 1 maps the comprehensive search strategy and rigorous selection process implemented to identify pertinent studies for this systematic review.

Following the initial stage of independent article screening using Rayyan ASReview software [14], the reviewers engaged in a consensus-building meeting to address any discrepancies that arose during the selection process. This collaborative approach fostered a balanced and objective evaluation of each study's relevance to the review objectives. To bolster confidence in the consistency of the screening process and minimise potential reviewer bias, measures of inter-rater reliability (IRR) were established. The primary indicator employed was the level of agreement between the reviewers, expressed as a percentage. This resulted in a high concordance rate of 90%, signifying a strong level of initial agreement. Furthermore, to account for the possibility of chance agreement, the reviewers calculated Cohen's kappa coefficient (κ). This more robust measure considers the agreement beyond what could be expected by coincidence. The resulting kappa coefficient of 0.85 indicated a substantial level of agreement between the reviewers, further solidifying the reliability of the screening process. These measures collectively demonstrate the rigour employed to curate a comprehensive and unbiased selection of studies for this systematic review.

2.3. Data Extraction and Coding

Aligned with the principles outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [11], data extraction aimed to systematically collect relevant information from the selected studies regarding data privacy considerations in postgraduate education. To ensure a standardised and efficient process, an open-source online tool called CADIMA [16] was utilised for data extraction. The data extraction form encompassed the following key elements:

- Study Characteristics: Author(s), year of publication, geographical location of the study (focusing on education programmes).
- Research Methodology: Research design employed (e.g., case studies, surveys), data collection methods (e.g., interviews with faculty or students), and approaches to data analysis.
- Sustainable Educational Data Analytics in Education: Specific educational data analytics techniques and tools used in the context of postgraduate education, along with details regarding how these practices addressed student privacy concerns.
- Postgraduate Education Context: Description of the postgraduate education setting (e.g., type of institution, programme focus), including any relevant aspects related to data governance frameworks or privacy policies.
- Data Privacy Considerations: Key findings related to the potential benefits and challenges associated with sustainable educational data analytics in postgraduate education, with a particular focus on student privacy risks and mitigation strategies explored by the studies.
- Research Limitations and Future Directions: Limitations identified by the study authors and potential areas for further research regarding data privacy and sustainable educational data analytics in postgraduate education.

Data coding within this systematic review was a manual process. It involved a thematic analysis approach, focusing on extracting and documenting core information such as author(s), study design, publisher, and key findings relevant to data privacy considerations in the context of postgraduate education. The extracted data were then systematically organised and presented in Table 1, which offers an overview of the data extraction and coding procedures employed in this review.

Table 1. Data Extraction and Coding.

Author(s)	Title	Source	Findings
[17]	Learning analytics in a shared-network educational environment: ethical issues and countermeasures.	International Journal of Advanced Computer Science and Applications (IJACSA)	Identified ethical concerns surrounding student data privacy in learning analytics, including potential for discrimination, lack of transparency, and student surveillance. Proposed countermeasures such as informed consent, data minimisation, and clear data governance practices.
[18]	Analytics in higher education: benefits, barriers, progress, and recommendations [Research Report].	EDUCAUSE Center for Applied Research	Highlighted the benefits of learning analytics for improving student learning and institutional effectiveness. Also identified are barriers such as faculty resistance, data privacy concerns, and lack of technical expertise.
[19]	A comprehensive AI policy education framework for university teaching and learning.	International Journal of Educational Technology in Higher Education	Proposed a framework for educating university faculty and students on responsible AI use in teaching and learning. This framework emphasises the importance of data privacy considerations within the context of AI-powered learning analytics.
[7]	Big data and the ethical implications of data privacy in higher education research.	Sustainability	Examined the ethical implications of big data analytics in higher education research, particularly regarding student data privacy. Highlighted the need for robust data security measures and adherence to ethical research principles.
[20]	Student perceptions of privacy principles for learning analytics.	Education Tech Research Dev	Investigated student attitudes towards privacy principles in learning analytics. Found that students generally support learning analytics but value transparency, control over their data, and clear communication about data usage.
[21]	A measurement of faculty views on the meaning and value of student privacy.	Journal of Computing in Higher Education	Explored faculty perceptions of student data privacy in the context of learning analytics. The study revealed a diversity of views, with some faculty members emphasising the importance of data privacy and others prioritising the potential benefits of learning analytics.
[22]	Technological barriers and incentives to learning analytics adoption in higher education: insights from users.	Journal of Computer-Assisted Learning	Identified technological challenges that hinder the adoption of learning analytics in higher education. The study also highlighted the need for clear data governance policies and faculty incentives to encourage the wider use of learning analytics tools.
[23]	Identification of 'at risk' students using learning analytics: The ethical dilemmas of intervention strategies in a higher education institution.	Educational Technology Research and Development	Examined the ethical considerations of using learning analytics to identify and support at-risk students. The study emphasised the importance of balancing early intervention with student privacy and avoiding potential student stigmatisation.

Table 1. *Cont.*

Author(s)	Title	Source	Findings
[24]	Privacy concerns and online learning of postgraduate students through the lens of stimulus–organism–response model.	Sustainability	Investigated the relationship between student privacy concerns and online learning engagement in postgraduate programmes. The study found that privacy concerns can negatively impact student engagement, highlighting the need for transparent data practices and strong privacy protections.
[25]	Students' privacy concerns in learning analytics: Model development.	British Journal of Educational Technology	Developed a model to understand the factors influencing student privacy concerns in learning analytics. The model suggests that factors such as perceived risk, trust in the institution, and awareness of data practices all play a role in shaping student privacy attitudes.
[26]	Revamping the academic library use data capabilities: the Greek library science postgraduates' perspective.	Library Hi Tech News	Explored postgraduate students' perspectives on data collection practices in academic libraries. The study identified concerns about student privacy and the need for clear communication about how library data are used.
[27]	Artificial intelligence education for radiographers, an evaluation of a UK postgraduate educational intervention using participatory action research: a pilot study.	Insights into Imaging	Evaluated a pilot study using AI-powered learning analytics in a postgraduate radiography program. While the study focused on the effectiveness of the intervention, it also acknowledged the importance of addressing student privacy concerns related to data collection and usage.
[28].	Supporting higher education with social networks: trust and privacy vs. perceived effectiveness.	Online Information Review	Investigated the relationship between trust, privacy concerns, and perceived effectiveness of social network-based learning analytics in higher education. The study found that student trust in the institution and the perceived benefits of learning analytics can mitigate privacy concerns, suggesting the importance of building trust and transparency around data practices.
[29]	The student expectations of learning analytics questionnaire.	Journal of Computer-Assisted Learning	Developed a questionnaire to measure student expectations regarding learning analytics. The questionnaire identified student concerns about data privacy and a desire for clear information about how their data will be used.

3. Results

This Section undertakes a critical synthesis of the literature review findings, specifically how they illuminate the established research objectives. Table 2 presents the germane findings from the literature review that pertain to the delineation of the types and dimensions of educational data typically collected in postgraduate programmes offered by higher education institutions (HEIs).

Leveraging the taxonomical framework established in Table 2, which dissects the multifaceted nature of postgraduate educational data, this study delves into the privacy concerns contingent upon the implementation of sustainable educational analytics within postgraduate programmes. The ramifications of this inquiry are documented in Table 3.

Table 2. Types and Dimensions of Postgraduate Educational Data.

Category	Dimensions	Scholarly Sources
Student demographics and background	Age, gender, nationality, prior academic background, work experience	[7,21,26–28]
Academic performance	Course grades, assignment performance, research project outcomes, thesis/dissertation quality	[18,21,23,28]
Engagement and participation	Class attendance, discussion participation, group project contributions, online forum activity	[20,23,25]
Learning outcomes	Standardised test scores (sustainability knowledge), applying sustainability principles, critical thinking (sustainability context)	[17,18,21–24,27,29]
Program satisfaction	Course content feedback, faculty effectiveness feedback, program structure feedback, overall satisfaction	[18–23,27,28]
Career outcomes	Employment data (job titles, sectors), career satisfaction (sustainability)	[20,21,24–26,28,29]

Table 3. Privacy Risks of Sustainable Educational Analytics in Postgraduate Programmes.

Risk Category	Description	Potential Consequences	Scholarly Sources
Data breach	Student data (demographics, performance, engagement, career outcomes) are exposed through accidental or malicious breaches.	Identity theft, discrimination, and reputational damage for students.	[7,17,18,20,22,26]
Misuse of data	Collected data are used for unintended purposes (commercialisation, profiling, disciplinary actions).	Student exploitation, and loss of trust in the institution.	[17,20,24,25]
Lack of transparency and consent	Students are unaware of data collection practices, usage, or sharing.	Mistrust between students and HEI, uninformed consent.	[20,21,24,25,29]
Algorithmic bias	Algorithms analysing student data perpetuate existing biases.	Unfair student assessments, and inaccurate program evaluations.	[21,22,24]
Chilling effect on participation	Fear of data misuse discourages participation in discussions, questions, and exploration of controversial topics.	Stifled critical thinking, and hindered learning environment.	[23,25]

The acknowledgement and mitigation of privacy risks through the implementation of appropriate safeguards is paramount. Such efforts will ensure the sustainable utilisation of educational analytics, thereby maximising the benefits for both students and postgraduate programmes. In this vein, the study conducted a comprehensive review of current practices and emerging frameworks designed to promote sustainable student data privacy in postgraduate programmes. These practices and frameworks are given in Table 4.

Table 4. Sustainable Student Data Privacy in Postgraduate Programmes.

Category	Practices	Description	Scholarly Sources
Current practices	Informed consent	Students understand what data are collected, how they are used, and who has access (builds trust and empowers choice).	[20,24,25,29]
	Data minimisation	HEIs collect only relevant data for specific analytics purposes (reduces exposure risk).	[18,22,23]
	Data anonymisation/pseudonymisation	Data are stripped of identifiers (anonymisation) or replaced with codes (pseudonymisation) for analysis.	[7,19,21]
	Secure storage and access controls	Robust security measures protect data storage and limit access to authorised personnel.	[17,21,26]
	Data retention policies	Clear policies specify how long data are stored before secure deletion (prevents unnecessary accumulation).	[18,23,27]
Emerging frameworks	Privacy-enhancing technologies (PETS)	Technologies like differential privacy (adding noise) and homomorphic encryption (analysing encrypted data) allow data analysis while protecting privacy.	[7,19,28]
	Data governance frameworks	Principles and practices are outlined for responsible data collection, use, and sharing (ensures compliance with regulations).	[17,18,24]
	Focus on student privacy rights	Frameworks emphasise student rights like access and control over their data (empowers students in data management).	[20,25,29]

4. Discussion

Table 2 provides an overview of the various types and dimensions of educational data typically collected in postgraduate programmes offered by HEIs. The table is accompanied by relevant scholarly sources that inform the understanding of each data category. The first category, 'student demographics and background', presents a demographic profile of the student body. This encompasses characteristics such as age, gender, nationality, prior academic background (e.g., undergraduate major), and work experience in given fields. Analysing these demographics allows HEIs to understand the programme's student composition and identify enrolment trends [20,21,25]. For instance, a programme might attract a higher proportion of students with backgrounds in environmental science. This information can be valuable for programme development and targeted recruitment efforts. Academic Performance data, another category in Table 2, focuses on measuring a student's achievements within the programme. It comprises course grades, performance on assignments, research project outcomes, and the quality of the thesis/dissertation. By analysing academic performance data, HEIs can identify strengths and weaknesses within the curriculum, as well as track individual student progress [18,21,23,28].

Table 2 also includes a category for 'engagement and participation'. These data capture a student's active involvement in the learning process, encompassing factors like class attendance, participation in discussions, contributions to group projects, and online forum activity. Analysing engagement and participation data allows HEIs to assess student engagement and identify areas where students might require additional support [20,23,25]. Learning Outcomes data, as presented in Table 2, assess the extent to which students have achieved the programme's core learning objectives. This might involve standardised tests measuring sustainability knowledge, evaluating a student's ability to apply sustainability principles to real-world problems, and assessing critical thinking skills within a sustainability context. Analysing these data helps determine the programme's effectiveness in achieving its learning goals [17,18,22]. 'Programme satisfaction' data gauge student feedback on various programme aspects. It includes feedback on course content, faculty effectiveness, programme structure, and overall satisfaction with the learning experience. Analysing programme satisfaction data allows HEIs to identify areas for improvement and enhance the student experience [18–20]. The final category in Table 2, Career Outcomes, tracks student career trajectories after graduation. This might include data on employment details (job titles, sectors), and self-reported career satisfaction. These data provide valuable insights into programme effectiveness in preparing graduates for careers [20,28]. Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

The utilisation of educational analytics in postgraduate programmes offers valuable insights into student learning and programme effectiveness. However, this data collection raises concerns regarding student privacy (Table 3). Educational institutions collect a comprehensive range of student data, encompassing demographics, academic performance, engagement metrics, and potentially even career outcomes. A data breach, whether accidental or malicious, could expose this sensitive information to unauthorised parties. This could have significant repercussions for students, including identity theft, job market discrimination, and reputational damage [24]. Student data collected for educational analytics could also be misused for unintended purposes. This encompasses commercialisation (e.g., targeting students with advertisements based on their programme or performance) or student profiling based on academic performance or engagement. Additionally, data could be used for disciplinary actions in ways students do not anticipate [17].

Furthermore, students might not be fully aware of the types of data being collected about them, how they are being used, or with whom they are shared. This lack of transparency can erode trust between students and the institution. Even if consent is obtained, it is crucial to ensure it is truly informed consent, meaning students understand the potential risks involved and can freely choose to participate [20,21,24,25,29]. Educational analytics often relies on algorithms to analyse student data. These algorithms can perpetuate existing biases, leading to unfair assessments of students or inaccurate programme evaluations. For example, an algorithm might unfairly penalise students from non-traditional backgrounds or learning styles [22]. Finally, if students fear their data will be used against them, they might be less likely to participate in online discussions, ask questions in class, or take risks in their learning (e.g., exploring controversial topics). This could stifle critical thinking and hinder a healthy learning environment [23].

Table 4 outlines a comprehensive set of practices and emerging frameworks employed by HEIs to ensure sustainable student data privacy within postgraduate programmes. Current practices focus on building trust and transparency with students. Informed consent, as supported by research from Ifenthaler and Schumacher [20], Majeed et al. [24], Mutimukwe et al. [25], and Whitelock-Wainwright et al. [29], empowers students by informing them about the data collected, their purpose, and who has access. This transparency fosters a sense of control and encourages student participation in data-driven initiatives. Data minimisation, highlighted by Bichsel [18], Klein et al. [22] and Lawson et al. [23], is another crucial practice. HEIs should collect only the data necessary for specific analytics objectives. This mitigates the risk of student exposure and ensures data utilisation aligns with its intended purpose. Data protection techniques such as anonymisation (removing identifiers) and pseudonymisation (replacing identifiers with codes) are also employed. Chan [19], Florea and Florea [7], and Jones et al. [21] discuss these methods, which permit data analysis while safeguarding student privacy. Robust security measures for data storage and access control are essential, as emphasised by Adejo and Connolly [17], Jones et al. [21], and Sant-Geronikolou and Kouis [26]. Limiting access to authorised personnel and implementing strong security protocols minimises the risk of data breaches. Finally, data retention policies, as discussed by Bichsel [18], Lawson et al. [23], and van de Venter et al. [27], specify data storage duration before secure deletion. This prevents unnecessary data accumulation and reduces privacy risks for students.

Looking beyond current practices, the table highlights emerging frameworks that further enhance student data privacy. Privacy-enhancing technologies (PETs) play a crucial role. Chan [19], Florea and Florea [7], and Wang et al. [28] discuss PETs like differential privacy and homomorphic encryption. These technologies allow data analysis while protecting sensitive information, offering a promising approach for future educational data analytics practices. Data governance frameworks, as highlighted by Adejo and Connolly [17], Bichsel [18], and Majeed et al. [24], outline principles and practices for responsible data collection, use, and sharing. These frameworks contribute to achieving compliance with relevant data privacy regulations (e.g., GDPR). Emerging frameworks also increasingly emphasise student privacy rights, including access to and control over their data [20,25,29]. This empowers students and fosters a culture of data ownership within the educational environment.

5. Conclusions

Educational data analytics (EDA) has emerged as a powerful tool for HEIs, enabling them to collect a wide range of data within postgraduate programmes. These data encompass student demographics and background, academic performance, engagement and participation, learning outcomes, programme satisfaction, and even career outcomes. However, for HEIs to achieve sustainability through educational data analytics, they must grapple with a range of student privacy risks. These risks include data breaches, misuse of student information, lack of transparency and informed consent, algorithmic bias, and a potential "chilling effect" on student participation due to privacy concerns. To mitigate these issues, HEIs can employ a multi-layered approach to ensure sustainable student data privacy in postgraduate programmes.

Building trust and reducing privacy risks can be achieved through informed consent procedures, data minimisation techniques (collecting only the essential data), and anonymisation or pseudonymisation of student data. Additionally, robust data security measures, such as secure storage, access controls, and data retention policies, are crucial for guaranteeing both physical and temporal data protection. Furthermore, the emergence of Privacy Enhancing Technologies (PETs) and data governance frameworks suggests a growing focus on both technological solutions and regulatory compliance within the context of student data privacy. However, striking a balance between maximising the benefits of data utilisation and protecting student privacy remains a significant challenge. HEIs must navigate this intricate balance by fostering open communication with students regarding data use, continuously evaluating and improving their data privacy measures, and staying abreast of evolving regulations. This ongoing effort is crucial for achieving long-term student data privacy within postgraduate programmes.

5.1. Limitations of the Current Analysis

This analysis prioritised the quality of evidence by focusing on peer-reviewed studies. While this approach ensured robustness, it inherently limited the scope of the review. Potentially valuable unpublished research (grey literature) on student data privacy practices in postgraduate programmes might have been excluded. Additionally, publication bias, where studies with statistically significant positive findings are more likely to be published, could have been introduced. However, the review mitigated these limitations to some extent by employing a comprehensive search strategy across multiple electronic databases to maximise the capture of relevant peer-reviewed studies. Furthermore, the review methods explicitly justified the rationale behind the inclusion/exclusion criteria, particularly the decision to exclude grey literature. The review was inherently limited by the specific inclusion/exclusion criteria used for selecting studies. This may restrict the generalisability of findings to all postgraduate education settings. However, by outlining the inclusion/exclusion criteria, the review allows readers to assess the applicability of the findings to their specific context.

5.2. Future Research Directions

The current review has established a solid foundation by analysing existing research on student data privacy in postgraduate programmes. Building on this knowledge base, future research can delve deeper into the practical application of educational data analytics while maintaining student privacy and fostering a sustainable approach to data use. In-depth case studies of educational data analytics implementation in diverse postgraduate programmes could offer valuable insights. These studies could examine the implementation process, focusing on how institutions balance data utility with student privacy while establishing sustainable practices. Exploring the effectiveness of specific techniques, identifying challenges encountered in real-world settings, and evaluating long-term data management strategies would provide practical guidance for other institutions seeking to achieve sustainable student data privacy. Quantitative research methods could shed light on the relationships between educational data analytics interventions, various student outcomes, and the long-term sustainability of data practices. Studies investigating correlations between educational data analytics initiatives and factors such as academic achievement, student engagement, programme satisfaction, and responsible data collection/storage practices would be valuable. This research could also explore the potential impact of educational data analytics on student trust and continued participation in data-driven initiatives.

Future research could delve deeper into the efficacy of specific techniques within the context of student privacy and data sustainability. Longitudinal studies could track the impact of a particular technique (e.g., learning analytics dashboards, social network analysis) on student learning over time, examining both academic outcomes and potential privacy concerns. Additionally, comparative studies could explore the relative effectiveness of different techniques for achieving specific learning objectives within different postgraduate programmes, while also considering their long-term data management requirements and impact on student trust. This would offer valuable insights into how institutions can optimise educational data analytics practices for diverse student cohorts and learning environments while prioritising sustainable student data privacy.

To gain a more comprehensive understanding of student data privacy and its sustainability, future research could utilise qualitative methods to explore the lived experiences of students and educators involved with educational data analytics initiatives. Conducting semi-structured interviews and focus groups with stakeholders would help capture their perspectives on the impact of educational data analytics on teaching, learning, the overall learning environment, and long-term data management practices. This would provide valuable insights into potential challenges, ethical considerations, and areas for improvement around sustainable data privacy practices. By pursuing these directions for future research, individuals can not only enhance the effectiveness of educational data analytics in postgraduate programmes but also ensure it is implemented in a way that safeguards student data privacy and fosters a culture of responsible data use for the long term.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/su16156377/s1.

Author Contributions: Conceptualization, P.N. and M.M.N.; methodology, P.N.; software, M.M.N.; validation, P.N. and M.M.N.; formal analysis, M.M.N.; investigation, M.M.N.; resources, P.N.; data curation, M.M.N.; writing—original draft preparation, M.M.N.; writing—review and editing, P.N.; visualization, M.M.N.; supervision, P.N.; project administration, P.N.; funding acquisition, P.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Foundation (SA), SRUG2205025721 and The APC was funded by the University of South Africa (Unisa).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: Acknowledge the support of two postdoctoral fellows at Unisa who assisted with checking the coding of data.

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

References

1. Le Blanc, D. A systems approach to the implementation of the Sustainable Development Goals. In *United Nations Department of Economic and Social Affairs (DESA), Report of the Open Working Group of the General Assembly on Sustainable Development Goals (A/68/970)*; United Nations: New York, NY, USA, 2014; pp. 1–26. Available online: https://digitallibrary.un.org/record/784147?ln=en (accessed on 18 March 2024).
2. United Nations. Sustainable Development Goals. 2015. Available online: https://sdgs.un.org/goals (accessed on 15 March 2024).
3. Tempelaar, D.T.; Rienties, B.A.; Giesbers, B. Stability and sensitivity of learning analytics based prediction models. In Proceedings of the 7th International Conference on Computer Supported Education, CSEDU, Lisbon, Portugal, 23–25 May 2015; pp. 156–166.
4. Mulder, I.; Jonges, F.M.; Rienties, B.A. Using learning analytics to save energy in education: A case study. *Sustainability* **2018**, *10*, 4372.
5. Lane, D. Privacy, big data, and the student: An ethical framework for educational data analytics. *J. Educ. Philos. Theory* **2017**, *49*, 104–127.
6. Mittelstadt, B.D.; Allo, P.; Taddeo, M.; Wachter, S.; Floridi, L. The ethics of algorithms: Mapping the debate. *Big Data Soc.* **2016**, *3*, 1–21. [CrossRef]
7. Florea, D.; Florea, S. Big data and the ethical implications of data privacy in higher education research. *Sustainability* **2020**, *12*, 8744. [CrossRef]
8. Prinsloo, P.; Kaliisa, R. Data privacy on the African continent: Opportunities, challenges and implications for learning analytics. *Br. J. Educ. Technol.* **2022**, *53*, 894–913. [CrossRef]
9. General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). Regulation (EU) 2016/679 of the European Parliament and of the Council. 2016. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679 (accessed on 1 March 2024).
10. Council of Europe. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data and Its Additional Protocol. 2018. Available online: https://rm.coe.int/1680078b37 (accessed on 10 April 2024).
11. 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. *J. Clin. Epidemiol.* **2021**, *134*, 178–189. [CrossRef] [PubMed]
12. Methley, A.M.; Campbell, S.; Chew-Graham, C.; McNally, R.; Cheraghi-Sohi, S. PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv. Res.* **2014**, *14*, 579. [CrossRef] [PubMed]
13. Rayyan. Faster Systematic Reviews. Available online: https://www.rayyan.ai/ (accessed on 10 April 2024).
14. ASReview. Join the Movement towards Fast, Open, and Transparent Systematic Reviews. Available online: https://asreview.nl/ (accessed on 10 April 2024).
15. Critical Appraisal Skills Programme (CASP). CASP Checklists. Available online: https://casp-uk.net/casp-tools-checklists/ (accessed on 10 March 2024).
16. CADIMA. Available online: https://www.cadima.info/index.php/site/about (accessed on 21 April 2024).
17. Adejo, O.; Connolly, T. Learning analytics in a shared-network educational environment: Ethical issues and countermeasures. *Int. J. Adv. Comp. Sc. Appl. IJACSA* **2017**, *8*, 22–29. [CrossRef]
18. Bichsel, J. *Analytics in Higher Education: Benefits, Barriers, Progress, And Recommendations [Research Report]*; EDUCAUSE Center for Applied Research: Boulder, CO, USA, 2012; Available online: http://www.educause.edu/ecar (accessed on 1 April 2024).
19. Chan, C.K.Y.A. comprehensive AI policy education framework for university teaching and learning. *Int. J. Educ. Technol. High Educ* **2023**, *20*, 38. [CrossRef]
20. Ifenthaler, D.; Schumacher, C. Student perceptions of privacy principles for learning analytics. *Educ. Tech. Res. Dev.* **2016**, *64*, 923–938. [CrossRef]
21. Jones, K.M.L.; VanScoy, A.; Bright, K.; Harding, A.; Vedak, S. A measurement of faculty views on the meaning and value of student privacy. *J. Comp. Higher Educ.* **2022**, *34*, 769–789. [CrossRef]
22. Klein, C.; Lester, J.; Rangwala, H.; Johri, A. Technological barriers and incentives to learning analytics adoption in higher education: Insights from users. *J. Comput. Assist. Learn.* **2019**, *31*, 604–625. [CrossRef]

23. Lawson, C.; Beer, C.; Rossi, D.; Moore, T.; Fleming, J. Identification of 'at risk' students using learning analytics: The ethical dilemmas of intervention strategies in a higher education institution. *Educ. Technol. Res. Dev.* **2016**, *64*, 957–968. [CrossRef]
24. Majeed, M.; Ghani, U.; Meng, W. Privacy concerns and online learning of postgraduate students through the lens of stimulus–organism–Response model. *Sustainability* **2022**, *14*, 11604. [CrossRef]
25. Mutimukwe, C.; Viberg, O.; Oberg, L.; Cerratto-Pargman, T.; Mutimukwe, C. Students' privacy concerns in learning analytics: Model development. *Br. J. Educ. Technol.* **2022**, *53*, 932–951. [CrossRef]
26. Sant-Geronikolou, S.; Kouis, D. Revamping the academic library use data capabilities: The Greek library science postgraduates' perspective. *Libr. Hi Tech. News* **2020**, *37*, 5–9. [CrossRef]
27. van de Venter, R.; Skelton, E.; Matthew, J.; Woznitza, N.; Tarroni, G.; Hirani, S.P.; Kumar, A.; Malik, R.; Malamateniou, C. Artificial intelligence education for radiographers, an evaluation of a UK postgraduate educational intervention using participatory action research: A pilot study. *Insights Into Imaging* **2023**, *14*, 25. [CrossRef] [PubMed]
28. Wang, W.; Lam, E.T.H.; Lung, M.M.; Chiu, D.K.W. Supporting higher education with social networks: Trust and privacy vs perceived effectiveness. *Online Inf. Rev.* **2021**, *45*, 207–219. [CrossRef]
29. Whitelock-Wainwright, A.; Gašević, D.; Tejeiro, R.; Tsai, Y.; Bennett, K. The student expectations of learning analytics questionnaire. *J. Comput. Assist. Learn.* **2019**, *35*, 633–666. [CrossRef]