

Received 21 December 2023, accepted 8 February 2024, date of publication 12 February 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3365368

TOPICAL REVIEW

Could the Use of AI in Higher Education Hinder Students With Disabilities? A Scoping Review

ORIANE PIERRÈS^{1,2}, MARKUS CHRISTEN², FELIX M. SCHMITT-KOOPMANN¹,
AND ALIREZA DARVISHY¹

¹Institute of Computer Science, Zurich University of Applied Sciences School of Engineering, 8400 Winterthur, Switzerland

²Institute of Biomedical Ethics and History of Medicine (IBME), University of Zurich, 8006 Zurich, Switzerland

Corresponding author: Oriane Pierrès (oriane.pierres@zhaw.ch)

This work was supported in part by the swissuniversities Project: P7—Accessible teaching in higher education, and in part by the Swissuniversities Project: P8—Swiss digital skills academy: accessible open educational resources.

ABSTRACT Literature reviews on artificial intelligence (AI) have focused on the different applications of AI in higher education, the AI techniques used, and the benefits/risks of the use of AI. One of the greatest potentials of AI is to personalize higher education to the needs of students and offer timely feedback. This could benefit students with disabilities tremendously if their needs are also considered in the development of new AI educational technologies (EdTech). However, current reviews have failed to address the perspective of students with disabilities, which prompts ethical concerns. For instance, AI could treat people with disabilities as outliers in the data and end up discriminating against them. For that reason, this systematic literature review raises the following two questions: To what extent are ethical concerns considered in articles presenting AI applications assessing students (with disabilities) in higher education? What are the potential risks of using AI that assess students with disabilities in higher education? This scoping review highlights the lack of ethical reflection on AI technologies and an absence of discussion and inclusion of people with disabilities. Moreover, it identifies eight risks associated with the use of AI EdTech for students with disabilities. The review concludes with suggestions on how to mitigate these potential risks. Specifically, it advocates for increased attention to ethics within the field, the involvement of people with disabilities in research and development, as well as careful adoption of AI EdTech in higher education.

INDEX TERMS Artificial intelligence, disabilities, higher education, educational technologies (EdTech), ethics, risk assessment.

I. INTRODUCTION

According to Russell and Norvig [1], the study of artificial intelligence (AI) refers to “the study and construction of agents that do the right thing” (p.22). What the “right thing” means depends on how the objective is defined (ibid.). In the context of higher education, several reviews have identified different applications of AI [2], [3], [4], [5], [6], [7], [8]. Zawacki-Richter et al. [7] distinguish four uses of AI: 1) profiling and prediction, 2) assessment and evaluation, 3) adaptive systems and personalisation, and 4) intelligent tutoring systems. Overall, the use of AI for tertiary education

is still experimental [4], and its adoption is at an early stage [9]. This is also reflected in the technologies used, as traditional AI techniques such as logistic regression or naive bayes classifiers are more widespread than more advanced methods such as neural networks and genetic algorithms [6].

Existing reviews also explore the benefits and risks of using AI in higher education. On the one hand, Ouyang et al. [6] argue that AI enables administrative staff and lecturers to take informed decisions based on predictions of student performance, learning status, or satisfaction. AI also provides students with learning recommendations, and it improves academic performance as well as online engagement and participation. These arguments are also supported by the conclusions drawn from the literature review by

The associate editor coordinating the review of this manuscript and approving it for publication was Meriel Huggard¹.

L. Chen et al. [10]. Similarly, Zawacki-Richter et al. [7] point out the possibility to facilitate admission decisions with predictions that let university employees focus on more complex cases. Moreover, the review by González-Calatayud et al. [3] suggests that students perform better with AI tutors than without.

On the other hand, the risks and ethical concerns of using AI are manifold. Some researchers stress the lack of pedagogical perspective, as the literature focuses on the technical aspects of developing AI applications [3], [6], [7], [11]. Zawacki-Richter et al. [7] emphasize the fact that technology is not a panacea, as well as the importance of supporting learners in a pedagogic sense. Furthermore, there are concerns about the lack of transparency of AI-based decisions [3], [6]. Some researchers also warn against biases in AI due to a lack of data diversity [6], [7], [12], [13]. There are also concerns about the privacy of students and security issues [7], [8], [12], [14]. Staff may also not be trained to use AI appropriately [3], [14]. Finally, students and higher education staff alike may not trust that AI-based proctoring systems make correct assessments [14].

Higher education institutions are seeking to make education more accessible, and AI could help in achieving this goal. In 2006, the signatories of the United Nations Conventions on the Rights of Persons with Disabilities pledged their responsibility in guaranteeing that people with disabilities have access to tertiary education. From a socio-medical perspective, a disability is the result of the “interaction between health conditions [. . .] and a range of environmental and personal factors” [15]. This definition elucidates that focusing on the health condition of individuals would be insufficient to improve accessibility. There is an increasing recognition that there is no one-size-fits-all approach to education. The Universal Design for Learning (UDL) approach recognizes that every learner has a unique way to understand information, to be motivated, and to express their knowledge [16]. Therefore, higher education should be more flexible and tailored to the diverse needs and abilities of students. AI provides the personalization and adaptability that could enable people with disabilities to study at the tertiary level. For instance, students with impairments may benefit from a flexible course schedule with adapted learning paces, as attending courses requires a lot of effort and concentration [17]. However, these benefits can only be achieved if the needs of people with disabilities are considered in the development and adoption of AI EdTech. As Heiman et al. [18] note, failing to account for accessibility needs during technology development can lead to increased costs and lengthier processes for people with disabilities.

Unfortunately, existing literature reviews fail to make a deeper analysis on the impact AI could have for students with disabilities. Very few papers mention people with impairments. While González-Calatayud et al. [3] briefly mention one study that considered students with disabilities for an adaptive learning platform with self-assessment in their

review, they provide no further analysis. Nigam et al. [14] caution that AI-based proctoring systems, i.e. systems that monitor students during examinations to prevent cheating, can increase students’ anxiety. Nevertheless, they do not consider how this may impact students with chronic anxiety or other mental health conditions. A recent systematic literature review found that knowledge in inclusiveness is limited and remains fairly new in learning analytics, as they could not find articles older than 2016 dealing with the topic [19]. This present study seeks to understand how the discussion of ethics and people with disabilities have been integrated in the recent literature. Specifically, it focuses on research articles presenting AI educational technologies (AI EdTech) that assess students to form or inform decisions in higher education. Assessment is defined broadly as “the process of gathering information and intervening in that information using some criteria in order to form a judgment” [20, p.4]. For instance, an AI could assess a student’s skills, interests, and preferences to recommend suitable universities [21].

Moreover, the field of AI EdTech needs to pay greater attention to ethics [2]. More specifically, discrimination is a serious concern in the use of AI [22]. According to Heinrichs [22], there are four aspects that characterise discrimination: 1) it is a situation when individuals are treated differently than others, 2) the treatment is, or at least is believed to be, disadvantageous towards the individuals, 3) the difference in treatment can be explained by the belonging to a specific group, and 4) the treatment is not in accordance with established weighting of ethical concerns. With this moralized definition, Heinrichs [22] stresses that to judge a decision as discriminatory, one needs to consider how it violates ethical considerations towards a group of individuals. According to Morris [23], there are seven ethical concerns for AI and accessibility: inclusivity, bias, privacy, error, expectation setting, unfeasible simulated data, and social acceptability. Firstly, inclusivity raises questions about the effectiveness of AI technologies “for diverse user populations” [23, p. 35]. For instance, text correction is less likely to work for people with dyslexia [24]. Secondly, bias is a source of harm in machine learning and can lead to discriminatory treatment of different groups of people [25]. This poses a problem when it affects the right to education, which is guaranteed by the United Nations and national legislation in countries like Switzerland. There are allocative harms, meaning that specific groups are systematically excluded from opportunities or resources that other groups receive, e.g. a group of people is given lower chances to be admitted at a university (ibid.). There are also representation harms, meaning that some groups are represented negatively or lack a positive representation (ibid.). For instance, natural language processing (NLP) models associate mental illness with negative terms like gun violence, homelessness, and drug addiction [26]. Furthermore, as Trewin [27] explains, it is challenging to ensure algorithmic fairness for people with disabilities since disabilities are diverse, can be multiple, and are

usually treated as outliers in the data used to train algorithms. Thirdly, people with disabilities are more subject to privacy issues because their disability can act as an identifying factor in an anonymised data set or the AI may deduce a disability from the data, leading to involuntary disclosure [23]. Fourthly, “error” refers to the fact that “many people with disabilities need to trust and rely on the output of an AI system without the ability to verify the output” (ibid., p.36). Fifthly, expectation setting issues arise when the capacity of AI is marketed as greater than what it can actually deliver, leading to false promises for people with disabilities who rely on these technologies (ibid.). Sixthly, simulated data are difficult to create for people with disabilities and often unrealistic (ibid.). Consequently, AI evaluations need to involve people with actual impairments (ibid.). Finally, “social acceptability” revolves around the question of whether a technology is better accepted because of the disability status of the user (ibid.). For example, Google Glass may not be accepted by the general public due to privacy concerns, but its use to help people with visual impairments may receive more acceptance [28]. These seven ethical concerns by Morris [23] are crucial to understand how AI use in higher education can potentially discriminate against students with disabilities. The European Commission [29] recommends involving individuals from vulnerable groups, such as people with disabilities, and investigating potential harm to ensure the development of trustworthy AI. Yet, there is not enough research on AI fairness that considers the perspective of people with disabilities [25], [27].

Taking the current literature into account, this study explores two main research questions. The first question aims to investigate the extent to which ethical concerns are addressed in articles that present AI applications in higher education – both in general and more specifically in relation to students with disabilities:

Research Question 1: To what extent are ethical concerns considered in articles presenting AI applications assessing students (with disabilities) in higher education? This requires answering two sub-questions:

- a) What are the ethical concerns mentioned in articles presenting AI applications assessing students in higher education?
- b) To what extent are students with disabilities mentioned?

The second main research question of this study focuses on assessing the risks of discrimination with AI applications presented in the literature. Special attention will be paid to the input variables used in AI applications, since data are often the source of AI biases [25]. By examining the input variables, it is possible to identify potential measurement bias, i.e. biases that arise due to measuring a concept with proxy data [30]. Furthermore, not all decisions are equally critical; there are low- and high-impact decisions [31]. High-impact decisions, such as hiring someone, can have a significant

impact on individuals’ lives and deprive them of their right to equal opportunities (ibid.). Low-impact decisions, such as a platform recommending a product to buy, are less consequential [32]. This distinction relates to allocative harms of biases. For that reason, the type of decisions taken by an AI in the articles will be extracted. Additionally, the degree of user control influences how algorithmic bias affects users as they can be empowered to reject or refuse the decision of the AI [32]. Hence, this review will examine who is making the decisions in AI applications.

Research Question 2: What are the potential discrimination risks in using AI that assesses students with disabilities in higher education?

- a) What are the input data of the AI EdTech?
- b) What type of decision is taken?
- c) Who is involved in the decision-making process and how?

Early results of this study were presented in a short paper at the European Workshop on Algorithmic Fairness 2023 [123]. This study contains the final results with a detailed discussion, as well as a description on the method.

This paper is structured as follows: first, the methodology of this scoping review, based on the PRISMA-ScR method, is described. Second, the results of the review are presented. In particular, this section reports on how many articles have discussed ethical concerns and the perspective of students with disabilities. The types of decisions made by AI, how humans are involved in the decision-making process, and types of variables used for the AI are also reported in this section. Third, the discussion section highlights the lack of ethical considerations and outlines eight discrimination risks for students with disabilities. Finally, the conclusion summarizes the answers to the research questions raised, discusses the limitations of the study, and provides suggestions for future research.

II. METHOD: PRISMA-ScR

To review the literature systematically, the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Review (PRISMA-ScR) method was followed. The PRISMA provides guidelines to guarantee the transparency, completeness and replicability of the review to the fullest possible extent [33]. The method includes an extension for scoping reviews, as they ask broader questions than systematic literature reviews and thus not all reporting items are relevant for scoping reviews [34]. This scoping review examines whether ethical concerns related to students with disabilities are considered during the development and use of AI in higher education. Then, looking at the description of AI in academic literature, this study assesses the potential risks these AI applications could pose for students with disabilities. In sum, this scoping review seeks to highlight any gaps in the literature regarding ethical considerations for students with disabilities in AI-based educational technologies.

TABLE 1. Eligibility criteria of the scoping review.

Type of criteria	Inclusion criteria	Exclusion criteria
Context	Applications only in higher education.	Applications not in higher education (e.g., applications in K-12 education). (EX1)
Focus of the article	<p>Presentation of AI-based systems (this could be an empirical test of an AI-based system or detailed presentation of an AI-based framework that required testing).</p> <p>Presentation of an AI that has a practical application in higher education, and is a complete system rather than just one element of an AI system.</p>	<p>Articles that do not present an AI-based system (e.g. general discussion of the potential of AI in higher education; survey of opinions on the use of AI; use of AI techniques to assess a case, but it is not an application). (EX2)</p> <p>Presentation of an AI-based system that does not explain how the results are used (e.g. identifies predictors and creates a prediction model, but does not say what the system does with the results), or the results do not have a practical application and the focus is on optimising the models (e.g. focus on the algorithm technique, on optimisation of algorithms). (EX3)</p>
Type of technology	Systems that use AI (machine learning, deep learning) to assess students and make decisions affecting students.	Rule-based systems (they are not AI) (EX4)
Type of application	AI technologies that take or inform decisions affecting students (e.g. admission into university, being flagged as at risk, getting graded, etc.).	<p>AI that does not assess students (e.g. it assesses teachers or university). (EX5)</p> <p>AI-based decision does not affect students directly (e.g. AI that measured student satisfaction to improve university performance). (EX6)</p>
Quality indicators	Article with a sufficient degree of detailedness on the presented system.	<p>Editorials, Conference reviews, Literature reviews, books, short paper for poster session or extended abstract, presentation of early projects (EX7)</p> <p>The article is not detailed (lack of details to present the AI-based system or lack of results because it is an early project) (EX8).</p>
Language	English.	Any other language than English. (EX9)
Time period	From 2018 to July 2022.	Any article published before 2018. (EX10)
Availability	Articles available in full through the university licenses of the ZHAW or the UZH.	Articles that must be bought or that cannot be found online. (EX11)
Uniqueness	Articles presenting a unique AI application.	Articles from the same authors that presented an AI application that was already included in the selection. (EX12)

A. PROTOCOL AND SCREENING REPORT

A protocol was written before the start of the review and updated throughout the research process by the primary author. The screening report gathers the articles that were screened in this review and indicates the exclusion criteria for articles excluded in the second and third rounds. The protocol and the screening report are available on Zenodo.

B. ELIGIBILITY CRITERIA

Articles were selected following a set of inclusion and exclusion criteria that were updated during the search and selection procedure (Table 1). Eligibility criteria focus on two aspects of the articles: content and form. Content criteria include “context”, “focus of the article”, “type of technology”, and “type of application”. Formal criteria comprise “Quality indicators”, “language”, “time period”, and “availability”.

The review included only studies in the higher education context, as students with disabilities face different barriers in higher education than in secondary education. For instance, while they still have the same needs as in secondary education, students with disabilities must self-advocate to receive accommodations in higher education [35]. Furthermore, to comprehend how current AI technologies could impact students with disabilities, this study required articles that focused on a comprehensive AI-based application. Publications that lacked clear use were deemed inadequate for assessing the potential future impact. Similarly, only AI-based systems that assessed and took or informed decisions affecting students were included.

Rule-based systems were excluded from the review since we focused on the more modern approach with machine learning techniques. Short contributions (e.g. poster sessions, extended abstracts), non-original work (conference reviews, literature reviews), editorials and books were excluded because they either did not provide enough details about an AI-based system or they did not present a new AI-based system. Moreover, only articles written in English were selected due to the large number of publications on this topic and the fact that most scientific research on AI is published in English. Further, articles published before 2018 were not included because the AI field is developing rapidly, making newer technologies more relevant to investigate. More recent articles are also more likely to present improved AI-based systems, making older publications less relevant. Additionally, a recent systematic review found that the topic of inclusiveness and students with disabilities is fairly new in the literature on learning analytics as they did not find any articles earlier than 2016 on the subject [19]. It can therefore be expected that recent articles can now account for this issue as it has been raised within the research community. Articles that could not be accessed freely with the licenses of the two universities of the authors were excluded because the two universities provide access to most journals and thus purchase costs could not be justified.

C. INFORMATION SOURCES, SEARCH, SELECTION OF COURSES OF EVIDENCE, FLOW CHART

Prior to searching for articles, the first author of the study performed an exploratory search of the literature in Web of Science and looked for existing systematic literature reviews. This preliminary phase helped define a search strategy and identify keywords. The final search syntax was the following (adaptations of the syntax due to database specifics are explained below); the three groups are coupled with the Boolean operator AND:

- Technology
 - “artificial intelligence” OR “machine learning” OR “deep learning” OR “natural language processing”
- Context

- “higher education” OR “tertiary education” OR “undergraduate education” OR “graduated education”
- Type of application
 - “assessment” OR “recommendation” OR “adaptive learning” OR “monitoring” OR “admission” OR “education technology” OR “tutor* OR “proctor*”

Then, the search and selection of articles was divided into three phases: 1) selection of articles in databases, 2) snowballing, and 3) additional review based on a reviewer’s feedback. The process is presented in the flow diagram (Fig. 1).

In the first selection phase, a total of 1,594 records were identified from four databases: Web of Science (1,089), Scopus (393), arXiv (12), and Google Scholar (100). For Web of Science, the parameters were the same except that the selected document types were “article”, “proceeding paper”, “book chapter”, and “data paper”. The search was performed from July 1-6, 2022. For Scopus, the parameters were the keywords from the search syntax + from 2018 to 2022 + English only + selected document type are article, conference paper, conference review. The search was performed from July 22-26, 2022. For arXiv, the search syntax was reduced because the first search retrieved only three results. The search syntax was (“Artificial intelligence” OR “Machine learning” OR “Deep learning”) AND (“higher education” OR “tertiary education” OR “undergraduate education” OR “graduated education”) AND (“assessment” OR “recommendation” OR “adaptive learning” OR “proctor”). The search was performed on July 20th, 2022. For Google Scholar, the search was done with the defined search syntax and from 2018 onwards. Only the first 100 results ordered by relevance were screened because they emerge out of a full-text search yielding much more potential results that cannot be further filtered. Google Scholar was still included in the process because it allows to check that articles deemed as relevant by the search engine have not been missed. The search was done on September 9th, 2022.

In the first screening round, the first author reviewed the titles and abstracts of 1,594 articles. Records were excluded when the titles and abstracts did not contain the keywords of the search syntax or when the keywords were used in a different meaning. For instance, deep learning can refer to an AI learning as well as a pedagogical concept of learning. Based on this, 1,252 records were excluded, and 79 duplicates were removed.

Then, the first author read the introductions and conclusions to further reduce the selection based on the eligibility criteria. This second screening round was necessary because the eligibility criteria were revised after the first screening, as the scope was initially too broad. At that stage, the availability of the text was checked as well.

In the third round of screening, the first and fourth authors independently reviewed 105 articles in full length.

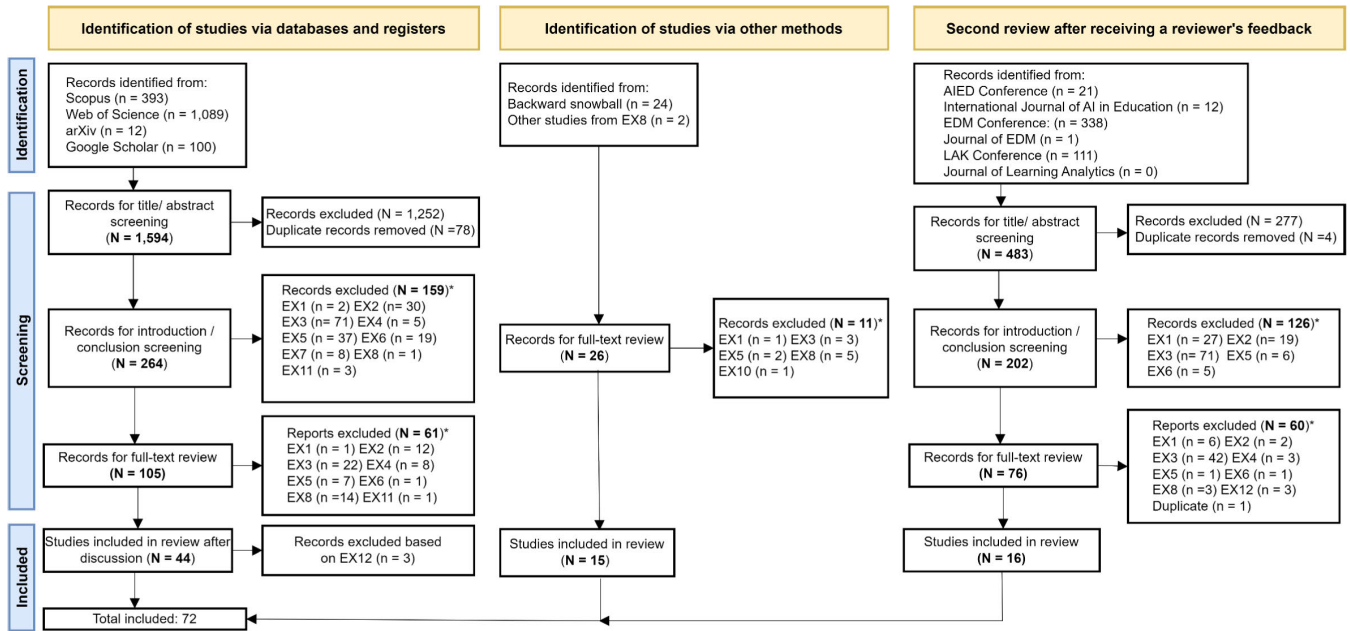


FIGURE 1. PRISMA flow diagram adapted from [33]. * = Total number of excluded reports does not equal the sum of all exclusion reasons because some articles had more than one reasons to be excluded.

Subsequently, they met to discuss potential disagreements and find a consensus. At the end of this stage, 44 articles were selected for analysis.

In the second selection phase, the first author followed the backward snowball technique to select relevant articles. Furthermore, to compensate the exclusion of papers due to a lack of details (EX8), the reviewer searched for publications that cited the excluded article and/or that were authored by the first, second, or last authors. The same eligibility criteria used in the first phase of the review were applied. 13 articles were identified with the snowball technique and two more recent articles were included based on the EX8 follow-up search.

During the analysis of the 59 articles, the first author noticed that two publication pairs were presenting the same AI application without significant changes to its design. To avoid inflating the results with multiple articles for one AI application, the articles that were less recent or less detailed were excluded (EX12).

After receiving feedback from a specialised academic journal, an additional literature review was performed to ensure that articles from important conferences and journals in the field of AI educational technologies were not missed. The following conferences and journals were reviewed: AIED conference, International Journal of AI in Education, Educational Data Mining (EDM) conference and journal, Learning Analytics and Knowledge (LAK) conference, and journal of learning analytics. The first author screened the articles following the same criteria as in the initial search. For the EDM conference, retrieving articles with keyword query was impossible which explained that they were not included in the first review; therefore, all articles from 2018 to 2022 were

retrieved manually. In total, 16 new articles were selected for analysis. Two articles were selected from the LAK conference that had not been captured by the query in the original databases. The other selected articles were published later than July 2022 which was the end date of the first review. One article was an extended version of an already selected article. Therefore, the shorter article was excluded following EX12. In total, 72 articles were analysed. The newly selected articles did not change the review's conclusions significantly. The only noticeable difference compared to the original selection of articles was a higher frequency of consideration for privacy and transparency issues.

D. DATA CHARTING PROCESS AND DATA ITEM

To extract data from the selected articles, a coding table (Table 2) was created and discussed among authors of the study. Data extraction was performed with one reviewer using MaxQDA. In case of uncertainty, a second reviewer was consulted to check the data extraction.

III. RESULTS

A. OVERVIEW OF THE ARTICLES

The 72 selected articles were sorted into four categories, as adapted from the work of Zawacki-Richter et al. [7]: 1) Profiling and predictions, 2) Assessment and evaluation, 3) Intelligent tutoring systems, and 4) Recommenders. Fig. 2 summarizes the distribution of the types of application in the sample of articles.

An interactive dashboard was created to let the readers filter the different articles that were coded for this analysis.

TABLE 2. Coding table of the review.

Data Name	Type of Data	Description
Application Types	Categorical: 1) profiling and prediction, 2) assessment and evaluation, 3) Recommenders, and 4) intelligent tutoring systems For each category, sub-categories were created inductively.	This code classified the type of AI-based applications presented. The code was based on the categories defined by Zawacki-Richter et al. (2019). The category of adaptive systems and personalisation was changed to recommenders because the former category could not be distinguished well from intelligent tutoring systems and profiling and predictions.
Were ethical concerns mentioned in the articles?	Binary: Yes or No	This code classified the articles whether the authors mentioned ethical concerns listed below.
Type of ethical concerns	Categorical: Inclusivity, Bias, Privacy, Error, Expectation Settings, Simulated Data, Social Acceptability Lack of transparency, lack of trust, lack of sustainability. (Multiple classifications possible)	This code classified the types of ethical concerns mentioned by the authors. The categories were defined deductively from the seven ethical concerns by Morris (2020). Three ethical concerns were added based on the ethical principles by Jobin et al. (2019).
Were people with disabilities mentioned in the articles?	Binary: Yes or No	This code classified the articles whether the authors mentioned people with disabilities.
Input Data	Categorical (dependent on application types) Multiple classifications possible as authors used multiple types of input data.	This code gathered all the input data (also called features or variables) used by the AI systems. Categories were created inductively.
Decision-maker	Categorical: AI, Student, Lecturer, Higher education staff.	This code indicates who was involved in the decision-making process. This code includes anyone cited as being informed or having a real power on the decision. Categories were created inductively.

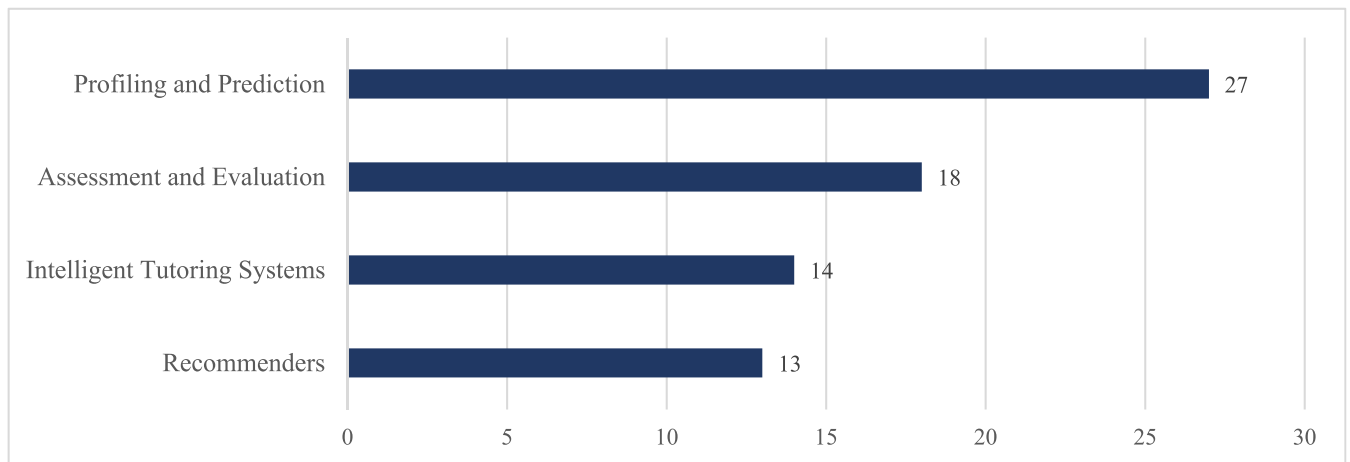


FIGURE 2. Distribution of application types in the 72 selected articles.

This dashboard also fosters accountability for the results reported in the following sections.

B. MENTIONS OF ETHICAL CONSIDERATIONS AND STUDENTS WITH DISABILITIES

Fig. 3 shows the distribution of ethical concerns mentioned in the 72 analysed articles. A little less than half of the

articles did not mention any ethical aspects. Among the 39 articles that mentioned ethical concerns, the majority considered privacy issues. Typically, authors indicated that data were anonymized or that privacy was guaranteed in accordance with regulations. Less frequently, authors explained the importance of guaranteeing privacy in the design of their applications. For instance, Robal et al. [36] argued that their

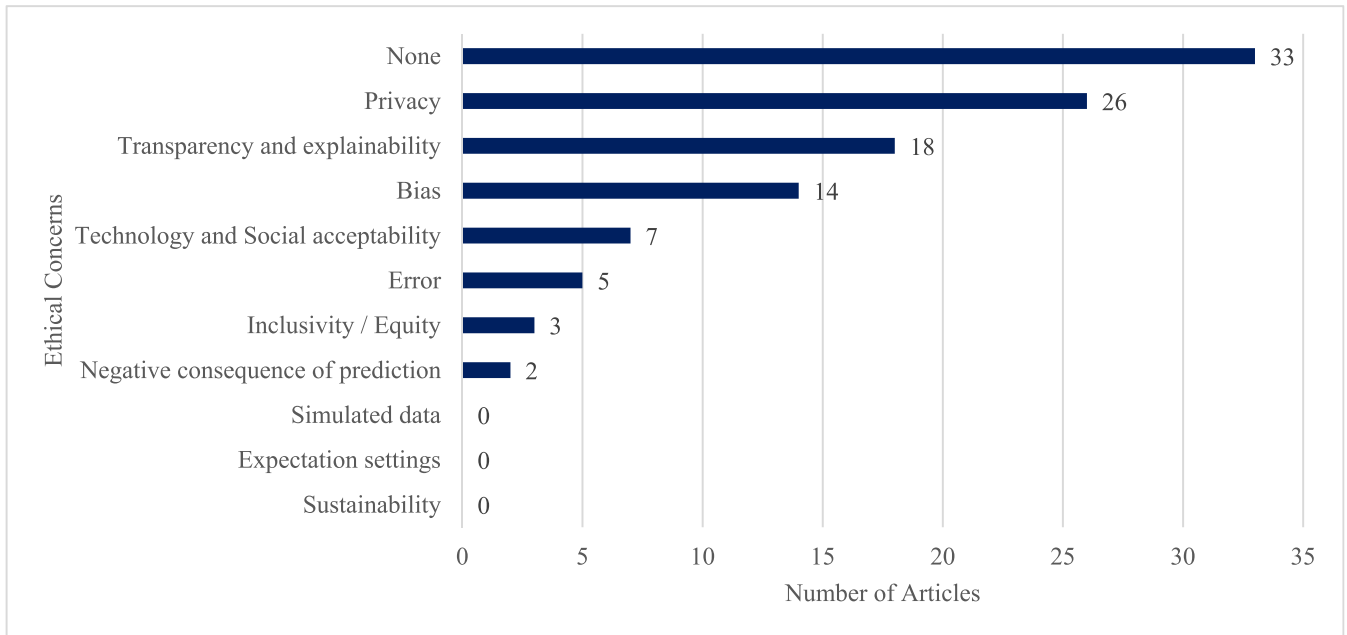


FIGURE 3. Distribution of ethical concerns mentioned in the 72 selected articles.

system for detecting attention loss in online learning was privacy-aware, as data were stored on local machines. Alike, Jia et al. [37] specifically investigated for the problem that algorithms may generate private content. Others specified that they did not include data such as gender due to privacy concerns (e.g. [38]).

The second-most commonly mentioned ethical concern was transparency and explainability. Several authors particularly emphasized the need to make predictions explainable (e.g., [39], [40], [41], [42], [43]). For instance, Ortigosa et al. [42] chose an algorithmic model that enabled them to explain which features contributed to a student being at risk of dropping out.

Finally, 14 articles mentioned considerations to mitigate biases. For instance, some excluded data like gender variables (e.g. [44]) or other demographic variables [45], while others checked that the dataset was not ill-balanced for variables known to create biases (e.g. [46]), or stated that models need to be trained periodically to reflect changes in the student population (e.g. [47]).

Only four of 72 articles mentioned students with disabilities [40], [48], [49], [50]. Among these four, Renzella et al. [48] and Nagy and Molontay [40] were the only ones to address or discuss accessibility considerations clearly. Renzella et al. [48] highlighted the importance of educators reaching out and engaging with students before introducing tutoring systems that may not be compatible with speech-generation technologies or that may increase anxiety by requiring students to record themselves. Nagy and Molontay [40] explicitly chose a colourblind-friendly colour palette for their application.

Xia [50] included disability as a variable to predict a final score and identify learning behaviour profiles that

need attention. Nevertheless, the author did not adequately define the variable and referred to students without disabilities as “learners with normal intelligence and good health” [50, p.9]. The author also suggested that learners with disabilities require a different type of intervention than the one for students without disabilities to support online learning. To make predictions on the probability of a student dropping out of their studies, Tsai et al. [49] included disability status in the variable “disadvantaged students”, which also comprised students from low-income households and students receiving financial support. Furthermore, the authors suggested engaging with and carefully listening to at-risk students and emphasized the importance of not leaving anyone behind.

C. POTENTIAL RISKS

This review extracted information on decision type, decision-maker, and input data that will feed a discussion on the potential risks of using AI in higher education for students with disabilities.

1) DECISION TYPE

This subsection describes the types of decisions that applications are affecting. For each kind of application, the decision types that are more likely to affect the studies of students are discussed in greater detail.

a: PROFILING AND PREDICTION

In profiling and prediction applications, three types of decisions were identified: 1) identification of students at risk of dropping out of a course, study programme, or school [20], [42], [46], [47], [49], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], 2) prediction of student performance prior to their admission [21], [64], [65], [66],

[67], 3) identification of students' learning profile [43], [50], [68]. The first two types of decisions have a more significant impact on access to studies than the last one.

Among the 72 articles, 18 presented an application to identify at-risk students. The output could be a probability to fail [49], [54], [58], [60], [63] or a categorical variable indicating that the student was at risk or not [52], [55], [59], [61], [20]. Additionally, several researchers used combined categorical variables with probabilities to present their outputs or added additional contextual information [42], [46], [47], [51], [56], [57], [59], [62]. This categorization of students is not trivial as supplementary information to understand the predictions was seldom provided. Sometimes, researchers discussed the importance of input features for predictions, but they did not include this information formally in the output, i.e. the decision-makers did not have access to the information [47], [61]. Van Petegem et al. [62] explicitly stated that teachers could use the feature weights to better understand why a student was predicted to pass or fail in a class. Comparatively, the solutions by Hussain et al. [59], Ortigosa et al. [42], Figueiredo and García-Peñalvo [57], and Eagle et al. [56] provided teachers with contextual information such as a dashboard visualising the online activities of students in the course or profile descriptions of students.

Six articles presented an AI-based system that predicted student performance before their admission. Three of the presented applications had a direct effect on admission selection: one provided a short-list of candidates for a graduate program based on their CV [69], another modified the weight of admission criteria to select students with the characteristics that predicted a higher performance at university [70], and the third one predicted enrolment of students to inform scholarship allocations [45]. The other three articles sought to inform admission decisions and strategic planning by highlighting skill or knowledge gaps and strengths [40], [71], [72]. Importantly, Nagy and Molontay [40] pointed out the self-fulfilling prophecy issue when students are predicted to fail in a program. As a result, the authors advised against focusing on the prediction and encouraged using predictors to understand how to increase the chance to graduate instead.

b: ASSESSMENT AND EVALUATION

Assessment and evaluation applications assessed 1) students' knowledge and skills [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], 2) their attention [36], [83], [84], [85], and 3) academic integrity and attendance [48], [86], [87], [88]. All solutions aimed to support instructors' grading and can impact student performance and chances of graduation. However, it should be noted that a majority of applications complemented instructors rather than replaced them; the final assessment was left to humans. Applications that assessed students' knowledge and skills usually classified students into categories rather than graded them. For instance, Rodriguez-Ruiz et al. [80] assessed communication skills of software engineers by categorizing students according to

their personality traits: verbally expressive, passionate, and perceptive. Similarly, systems assessing students' attention usually indicated to lecturers what a student was doing (e.g., listening, looking at phone, taking notes, chatting) [36], [83], [84], or gave an attention score [85]. Alike, the solutions provided instructors with binary output that checked academic integrity [48], [86], [87] or class attendance [88]. However, the application by Robal et al. [36] recognized student's attention to a video lecture which then generated one of the three automatic alerts: automatic pausing of a video, visual alert, or auditory alert.

c: RECOMMENDERS

Eight articles introduced recommenders for learning materials [38], [89], [90], courses [91], [92], topics to review [39], forum posts [93], or learning paths [94]. This type of decision has a rather low impact as it simply provides a list of courses or materials that could be of interest to students. In comparison, five articles utilised AI to recommend higher education institutions or programmes to students [21], [64], [65], [66], [67]. These recommendations can potentially impact students' decision to apply to a specific programme or university, impacting their future job and salaries. These applications provide a list of recommendations. Some also added a percentage of success to the recommendations [65], [67].

d: INTELLIGENT TUTORING SYSTEMS

Intelligent Tutoring Systems were divided between 1) those with a conversational agent [95], [96], [97] that usually sought to interact with students to enhance learning or provide support and 2) those without a conversational agent [37], [41], [43], [44], [98], [99], [100], [101], [102], [103], [104], [105] that typically provided feedback to programming submission, written works, or technical sketches. Interestingly, Azcona et al. [44] introduced an AI tutor to assist programming courses through several means: predicting grades and issuing alerts to students, offering examples of successful student work to encourage improvement, and informing students that if they had any doubts, they could contact the lecturer.

2) DECISION-MAKER

This sub-section outlines how various stakeholders (students, lecturers, faculty staff) were involved in the AI-assisted decision-making process. Note that their involvement does not imply final decision-making authority; rather, they receive information from AI outputs to inform decisions. Table 3 gathers the number of articles that involved each different stakeholder.

Across all applications, students and lecturers were equally frequently involved in the decisions informed by AI. However, the frequency diverged among application types. In profiling and predictions, the decision-makers were typically the lecturers or faculty staff who received information from the AI system and then decided on how to intervene or grade.

TABLE 3. Involvement of AI, student, lecturer, faculty in 72 articles. Cells in darker blue indicate the human stakeholder that was most involved in the respective application types.

Application type	Student	Lecturer	Faculty Staff	Unclear	Total articles
Profiling and Predictions	10	16	13	4	27
Assessment and Evaluation	6	13	5	1	18
Recommenders	12	6	0	0	13
Intelligent Tutoring Systems	13	6	0	0	14
Total Decision makers	41	41	18	5	72

For instance, for at-risk predictions, the most commonly proposed intervention was for the lecturer to send an email to the identified at-risk students, as seen in the works by Olivé et al. [61], Ciolacu et al. [52], or Le et al. [60]. Students were usually informed about AI outputs, but rarely had a more active role in the process. For instance, Andrews-Todd et al. [73], Datta et al. [95], and Wang and Chen [82] presented applications that informed students about AI-based assessment, without getting further control on their evaluation.

Comparatively, some applications let students decide whether to use the application. For instance, Ciolacu et al. [52] highlighted that students voluntarily agreed to receive at-risk predictions. Hellings and Haelermans [58] made students the sole decision-maker by providing them with a learning analytic dashboard featuring their predicted grade. In [49], the students were not informed about the prediction, and therefore they were not technically involved. Their AI application provided a list of students with a high risk of academic failure to academic counsellors, who then informed teachers. However, academic counsellors and teachers were encouraged not to label students, to engage in a dialogue with students, to listen to them, and to let them express concerns before deciding on an intervention. Additionally, the attention assessment system by Robal et al. [36] stood out by developing a tool that allowed students to self-monitor and choose not to use it. In comparison, the three other attention assessment applications [83], [84], [85] aimed to inform lecturers or faculty staff about students' attention to engage them.

Furthermore, students were more often involved in recommenders and intelligent tutoring systems. When it comes to recommenders, students can decide whether to follow the recommendations. However, only Hur et al. [39] and Morsomme and Alferez [92] accompanied their recommendation with an explanation to aid student comprehension. Intelligent tutoring systems generally aimed at enabling students' independent learning by providing feedback and recommendations. For

instance, both Kim et al. [99] and Vytasek et al. [105] presented an app to give feedback on the writing quality of students' essays before submission. Similarly, the conversational agent presented by Rossi et al. [96] suggested learning materials to enhance students' comprehension, but if unsatisfactory, the student could ask for help from a tutor. The language tutor by Schlippe and Sawatzki [97] not only graded and gave feedback to students, but also utilised peer-review to support them, thus granting students some form of control over correct answers. Still, in several articles, lecturers could also oversee the students' progress [44], [95], [96], [101].

The involvement of AI was generally limited to providing additional information to lecturers, faculty staff or students. For example, Rodriguez-Ruiz et al. [80] emphasized their tool was not for automatic grading but was complementary to the lecturers' assessment. Some also included human oversight to address inaccuracy or potential AI errors. For instance, Renzella et al. [48] presented an AI that checked students' real identity before taking an online exam and emphasized the importance of involving tutors in reviewing the alerts to avoid any negative impacts on students. Similarly, Sunaryono et al. [88] involved students in an AI-based attendance system by sending a confirmation that their attendance has been successfully recorded. However, there were a few instances where the AI took on a greater decision-making role. For instance, in [69], the faculty staff delegated the selection of candidates to the AI and only received a list of "good" candidates. Furthermore, two articles presented a knowledge and skill assessment tool that worked completely automatically and adjusted itself to students [75], [81].

3) INPUT DATA

This sub-section describes the data types used to train AI models. Table 4 summarizes the distribution of input data types in each application category.

TABLE 4. Distribution of input data types in AI-based applications. Cells highlighted in darker blue indicate the data type was mostly used for the respective application types.

Application type	Interaction log	Student socio-demographics	University & professional qualification	Collected preferences, interest, motivation, & appreciation	Text data	Audio data	Image & video data	Total articles
Profiling and Predictions	17	15	20	2	0	0	0	27
Assessment and Evaluation	5	3	3	0	7	1	10	18
Recommenders	6	3	10	3	1	0	0	13
Intelligent Tutoring Systems	3	1	1	1	11	1	1	14
Total All Applications	31	22	34	6	19	2	11	72

Data representing university and professional qualification and interaction log data were the most widely utilised across all applications. University and professional qualification data included information that was traditionally collected by universities such as grades, the type of courses that were passed or failed, or prior school or university. This data type was particularly employed to identify at-risk students and recommend learning materials, courses or programs. For example, Tsai et al. [49] used academic performance, student loan application, number of absences from school, and number of alerted subjects in the first and second semesters to identify students at risk of dropping out.

Digitalization and online learning development enable universities to leverage more information on students' learning behaviour. The interaction logs encompassed variables related to the activities and features of an online learning platform, such as the number of views to an online activity or a forum, or the time spent learning on a platform, but also more active forms of actions on an online platform like the student's number of attempted quizzes, number of submissions, number of forum posts, or number of emails sent and received. This information was mostly employed in profiling and prediction applications. Furthermore, socio-demographic information (e.g., age, gender, hobbies, parents' occupation, financial aid) was often used with interaction log data or university and professional qualification to predict students' academic performance or interest in a programme. For instance, Obeid et al. [21] asked students to enter personal information such as gender, living country, preferred

language, favourite hobby, role model in the family, parents' work domain as well as favourite/ least favourite subjects and previous school to recommend suitable universities to students.

Authors developing intelligent tutoring systems as well as assessment and evaluation tools usually took advantage of natural language techniques to analyse textual data and provide students with feedback. Typically, they assessed programming scripts [41], [76], [78], [102], students' essay or written report [37], [74], [79], [98], [101], [105], or chat or forum interactions [73], [95], [96], [97].

The use of video data was largely to assess students' attention, employing facial and eye recognition techniques, but also to assess students' comprehension as seen in the work by Holmes et al. [75] and Shobana and Kumar [81].

IV. DISCUSSION

A. TO WHAT EXTENT ARE ETHICAL CONCERNS CONSIDERED IN ARTICLES PRESENTING AI APPLICATIONS ASSESSING STUDENTS IN HIGHER EDUCATION?

In the literature presenting AI applications assessing students in higher education, ethical considerations are not systematic and remain superficial. This finding is in line with the findings of Zawacki-Richter et al. [7], who highlighted the lack of reflection of challenges and risks in AI educational technologies and a lack of pedagogical reflection. If known ethical concerns are not addressed in the literature, AI EdTech risks impacting students negatively. Privacy is the most widely known and addressed ethical concern in the literature, which is probably due to the importance of data protection laws in

recent years. However, privacy concerns are not consistently and deeply considered in the literature. This concurs with the findings of Khalil et al. [19] in the literature on learning analytics. Transparency and explainability are mentioned only in a quarter of the 72 selected articles. Similarly, bias is considered only in about 20% of the articles. Greater attention to these ethical aspects is essential because bias can lead to erroneous classification or even exclusion from a system. Furthermore, transparency can manage expectations and hold AIs accountable for their outcome. Explainable AI can also help users understand the AI outputs, facilitating decision-making.

This scoping review also highlights the missing perspective of students with disabilities. AI EdTech holds the promise to personalize higher education, but if designed without considering students with disabilities, EdTech may end up only serving the traditional students rather than the students universities are trying to support [106]. Moreover, there were no AI evaluations to measure the impact on students with disabilities. Thus, research presenting AI prototypes that assess students fail to account for the potential impact for students with disabilities. This is all the more problematic as students with disabilities are often underrepresented due to the difficulty in disclosing one's disability [107], and underrepresentation may lead to misclassification [108].

These findings suggest that while ethics and inclusion are discussed in the general field of AI Edtech (see e.g. [19], [108], [109], [110]), reflections on these topics do not seem to be integrated in the work of researchers who develop new applications. Creating inclusive and responsible technologies from the start is most beneficial because it takes more time, effort, and money to remediate inaccessible technologies [18]. To address this mismatch, Selwyn [110] suggested including ethics as a mandatory subject in the curricula of data scientists and those in charge of procurement decisions in higher education institutions. Researchers are also invited to report their efforts to address ethical concerns as well as limitations of their AI-based application regarding people with disabilities.

B. WHAT ARE THE POTENTIAL RISKS FOR PEOPLE WITH DISABILITIES TO USE AI IN HIGHER EDUCATION?

In total, eight potential risks of using AI in higher education for students with disabilities were identified:

- Risks associated with the choice of data:
 1. Interaction log data
 2. Background information: hobbies and financial aid
 3. Text data
- Risks associated with the decision type:
 4. Simplistic classification
 5. Recommendations that do not consider accessibility needs
- Risks associated with the involvement of stakeholders
 6. Monitoring students' faces

7. AI as a decision-maker
8. Low student involvement

Each potential risk is discussed in the following sub-sections.

1) RISKS ASSOCIATED WITH THE CHOICE OF DATA

First, many researchers sought to take advantage of the availability of interaction log data. The advantage of this type of data is that it captures what students are doing rather than who they are. However, even that type of data is not free from potential bias. The accessibility of online platforms may influence interaction logs of students with disabilities [106]. Interaction logs often included the number of times activities were viewed or the time spent on the platform. Barriers on online platforms such as illogical structure, improper labelling, low colour contrast, lack of alternative text for visual information, or lack of keyboard navigation support may hinder or even completely prevent students with disabilities from using the platform. This might be particularly relevant for students with a visual impairment who use screen-readers or magnification software and for students with a mobility impairment who navigate online with their keyboards or other alternative systems. Moreover, an attention deficit and hyperactivity disorder (ADHD) could potentially influence the number of interactions with the online platforms, as ADHD students might need to pause and review materials more often. It is however not clear whether or how ADHD students might have different interaction logs than students without ADHD, as no research on this topic could be found.

Second, the use of background information such as hobbies or financial aid can potentially be discriminating for students with disabilities. Many hobbies are not accessible to people with disabilities. For instance, Steinhardt et al. [111] identified several barriers for children's participation in hobbies such as the children's own preferences and competences, their social skills, the lack of adapted activities (especially in rural areas), or negative attitudes among peers. These barriers can affect people with any type of disabilities, even if they are likely to be affected differently. Financial aid may also be associated with disability status. Students with disabilities often have lower income [112]. During the COVID-19 pandemic, students with disabilities in the United States were more likely to experience greater financial hardship [113]. Many countries also provide financial aids to students with disabilities to facilitate access to higher education because of the higher costs of living with a disability [114]. Hobbies and financial aids were used in applications for admission (recommendation and performance predictions) and at-risk predictions. Especially in the case of admission, researchers must check that these variables do not hinder opportunities of students with disabilities. This is crucial because incorrect recommendations could create a feedback loop that perpetuates a situation where the same groups of students are consistently funnelled into certain universities.

Third, text data were often used to assess students' knowledge. Natural language processing techniques may be unable to analyse or retrieve information from texts written by students with cognitive or intellectual disabilities who, for instance, may misspell words due to dyslexia [24]. Consequently, students with disabilities may end up becoming excluded from systems that automatically analyse texts, and feedback may be more prone to errors. Moreover, Guo et al. [24] argued that "conversational agents may not work well for people with cognitive and/or intellectual disabilities, resulting in poor user experience" (p.4). They advised training the AI with diverse data to overcome this issue. However, considering that the perspective of students with disabilities is largely ignored in the literature, conversational agents are unlikely to be adapted to the needs of students with disabilities. Guo et al. [24] also suggested providing different modes of communication such as sign language, pictures, or icons, but this possibility was never presented in the selected articles.

2) RISKS ASSOCIATED WITH THE TYPE OF DECISION

Fourth, risk predictions often relied on numerical and categorical classifications which are simplistic and lack information on why students are at risk. Students with disabilities may require specific interventions. For instance, if the problem is that the platform is not accessible to a screen-reader user, then the intervention should be making this platform accessible to the student. Likewise, if a student with dyslexia struggles to demonstrate their knowledge through written exams, modifications to the exam format should be made. This undermines the claim that AI personalizes higher education. One solution is to provide results that may explain why a student is at risk and what could be improved. For instance, Van Petegem et al. [62] specifically chose an algorithmic model that allowed for interpretability of the at-risk predictions. Additionally, intervention systems can also be centred on talking and listening to students to understand them better, as was the case in [42] and in [49]. Another solution is to investigate whether some modules could have accessibility issues by comparing dropout rate and feedbacks between disabled and non-disabled students as suggested by Cooper et al. [109]. Still, although their method sounds promising, it raises privacy concerns as students may not necessarily be willing to disclose their disability to inform predictions on their performance.

Furthermore, students with disabilities may systematically end up in the at-risk category. A study on the OULAD dataset that gathered student data from a virtual learning environment found that students with disabilities were under-represented, which led in some cases to a higher probability to be classified as at-risk [108]. One may argue that it is not problematic that students are put into the "at risk" category, as these students then receive more attention from their lecturers. However, students with disabilities are often stigmatized and expected to perform poorly [35], [115], [116].

Thus, these categorisations may reinforce existing biases. Conversely, it is also possible that it could ease the problem of self-advocacy in higher education, as students with disabilities must disclose their disability and ask for specific accommodations to study. In particular, many students with invisible disabilities such as ADHD, chronic fatigue, or autism find disclosure emotionally difficult [35], [115], [116]. Consequently, more research is needed to understand whether students with disabilities are likely to be classified as at-risk. Additionally, researchers should seek the opinion of students with disabilities on these systems. These students may be willing to accept an erroneous classification as "at-risk" if it results in teachers becoming more proactive in addressing the specific needs of their students.

Fifth, the recommendation systems used for school applications and learning materials may fall short in promoting inclusivity and providing personalized recommendations as advertised. Despite their ability to suggest schools or courses that align with a student's academic performance, these systems lack personalized information on accessibility. To enhance inclusivity, the recommendation systems could provide details on the availability of disability offices and overall university accessibility. Similarly, for course and learning path recommendations, it would be helpful to thoroughly consider the accessibility of learning resources. Nonetheless, further research is needed to identify the recommendation features that students with disabilities would find useful.

3) RISKS ASSOCIATED WITH THE INVOLVEMENT OF STAKEHOLDERS IN THE DECISION-MAKING

Sixth, attention assessment raises concerns for continuous surveillance [14]. This holds true for all students, but especially for students with anxiety, who may feel under pressure, which increases barriers to succeed in higher education. Students with ADHD may also be systematically flagged and thus penalized. In that case, it is critical to let students decide whether they accept to be monitored or not like in the case presented by Robal et al. [36]. Moreover, the use of facial recognition may not work for students with unusual face features or students with visual impairments, e.g., if they wear sunglasses [24]. Moreover, Raji et al. [117] warned against AI that are not functioning because it is conceptually not possible to make inferences based on certain inputs. The authors gave the example of AI using physical appearance to infer personality traits [117]. While the emotion recognition was not really employed in the literature, two cases [75], [81] applied face and eye recognition to infer students' real knowledge. It is, however, questionable whether there is a founded causal link between someone's facial appearance and knowledge.

Seventh, at the moment, AI EdTech is designed to support and aid lecturers or faculty staff rather than replace them. However, it is conceivable that future developments in AI could delegate more decision-making responsibilities

to technologies. For example, currently, predictions regarding students who may be at risk are only meant to inform humans, who ultimately decide on the necessary interventions. Nonetheless, there are indications that automation may be possible in the future, such as through the sending of automated emails. Efforts to develop autonomous systems are also evident in the creation of intelligent tutoring systems. These systems offer students the opportunity to learn at their own pace online, which is particularly beneficial for students with disabilities since it helps to eliminate barriers [17]. Nonetheless, as the involvement of humans is reduced, students with disabilities may not have the possibility to request course adaptations. Furthermore, they usually cannot provide feedback to AI EdTech, i.e. the system cannot learn from them [107]. Higher education institutions need to maintain a certain level of flexibility to make reasonable accommodations. Still, it should be remembered that even when humans are involved, AI EdTech needs to be regularly reviewed to avoid automation bias, i.e. the human bias that tends to believe AI output [118]. Considering the lack of staff training [3], [14], inclusion training is crucial to ensure the rights of students with disabilities pursuing higher education.

Finally, the involvement of students was in most cases limited to receiving information on chances of success and learning progress. While this may be because lecturers and university staff are the traditional decision-makers in higher education, empowering students with information and decision-making power could be most beneficial. AI can be employed to assist students rather than control them. For instance, Robal et al. [36] presented a tool for students to detect their attention loss during video lectures, which they could choose to disable. While this has not been tested, their tool could be an assistive technology for students with ADHD. An assistive technology is defined as “any item, piece of equipment, or product system, whether acquired commercially off the shelf, modified, or customized, that is used to increase, maintain, or improve functional capabilities of a [person] with a disability” [119]. Technologies that may be considered optional or a luxury for students without disabilities can prove to be highly beneficial for students with disabilities. A prime example of this is the use of writing assistance tools like Grammarly, which can enable students with learning disabilities to write creatively. In the literature, only Kim et al. [99] presented a similar tool to improve argumentation quality. Focusing on developing AI-based educational technology that assists students, rather than just instructors, may be a crucial factor in advancing inclusion in higher education.

V. CONCLUSION

There is a clear lack of ethical consideration for students with disabilities in articles presenting AI EdTech that assesses students in higher education. A little less than half of the selected 72 articles did not address ethics at all, and those that did focused mainly on privacy, transparency/explainability,

or bias. Furthermore, the perspective of people with disabilities is largely missing.

This scoping review identified eight discrimination risks associated with the use of AI EdTech for students with disabilities, particularly emphasizing the potential for bias and exclusion. This raises concerns about the adoption of AI EdTech in higher education as it may hinder efforts towards greater accessibility and inclusion in this sector. It is important to note that the identified risks do not imply that AI EdTech should not be used in higher education to inform or take decisions affecting students. Nevertheless, it is crucial that AI EdTech are developed with accessibility and ethical concerns in mind to ensure that they benefit everyone. Additionally, human oversight and active efforts to incorporate the perspectives of students with disabilities are necessary to address these ethical concerns and promote greater accessibility and inclusivity.

A. THEORETICAL IMPLICATIONS

This study contributed to the AI Ethics field by identifying risks that are specific to the use of AI in higher education for students with disabilities. This publication aimed at sparking a discussion within the research community to consider the diversity of students and their subsequent perspectives. While this article focused on students with disabilities, it should be noted that the risks may also impact other groups such as students from a lower socio-economical background, different genders, or diverse ethnicities. For instance, the input variable “hobby” can also correlate with gender and socio-economical background and thus raise similar issues as for disability. Accounting for these risks is therefore crucial for more than students with disabilities. Future theoretical research could tackle an intersectional approach. For instance, it could investigate how disability status relates to gender and race in higher education and how it affects degree completion.

B. EMPIRICAL IMPLICATIONS

Empirical research is needed to test to what extent the identified risks impact students with disabilities. Several research questions can be raised, for instance whether and how interaction log data of students with disabilities (e.g., ADHD, visual impairment, and mobility impairment) differ from students without disabilities, and to what extent this affects predictions. Future research should also investigate the efficiency of solutions to mitigate risks. Qualitative research could also gather the opinions of students with disabilities to understand their preferences and risk tolerance if AI reduces barriers in higher education. Students with disabilities could also provide valuable information on features that could ease their access to higher education degrees.

C. PRACTICAL IMPLICATIONS

This review is a call for researchers developing AI EdTech to integrate ethical considerations in their research. These can be included when justifying the choice of the technological

design or at the end of articles when discussing future research and potential implementation. Doing so not only signals that risks have been taken into consideration, but also encourages the community to develop responsible AI EdTech.

Developers and public procurers are invited to keep in mind the commitment of higher education institutions for greater inclusion before acquiring or buying AI EdTech. EdTech that addresses and mitigates ethical concerns is to be favoured. For instance, systems using text data should include texts written by students with cognitive and intellectual disabilities, and conversational agents need to provide different modes of communications [24]. Moreover, practitioners need to keep in mind that technology is not always the solution to mitigate risks. Education is a complex phenomenon, which means an effective intervention is not necessarily the most efficient one in terms of resource management [120]. To avail of technology, public institutions need to adopt a holistic vision that utilizes both human and technological strengths alike [120].

Additionally, students with and without disabilities may benefit more from AI EdTech that assists them rather than controls or monitors them. Developers and public procurers are therefore encouraged to pursue this goal. The release of ChatGPT opened many opportunities to develop personalised assistive AI EdTech.

D. LIMITATIONS

This scoping review has several limitations. First, for most of the review process, there was only one reviewer. To reduce selection bias, a second reviewer was involved in the third screening of the articles, and the two reviewers discussed issues until consensus was found. In the other steps, if the main reviewer had any doubts, the second reviewer was consulted. Second, the review only includes articles written in English. This language restriction was established due to the quantity of existing papers and the fact that most scientific research on AI is published in English. Third, some researchers published their progress on an AI-based system in different papers. It is possible that some authors published their works in separate papers and therefore, the part on ethics is included in a different paper or that further tests were performed. However, this limitation is mitigated by the fact that papers with few details were excluded from the analysis, and we looked for newer and detailed research. Moreover, during the screening, it did not seem that this was a recurring issue. Fourth, the analysis could not differentiate systematically for each type of disabilities due to a lack of empirical analysis on the subject and the diversity of disabilities. Still, this study emphasized as much as possible the types of disabilities that could be affected by the risks raised in the analysis. Finally, this study has focused on articles presenting AI applications that assess students and inform or take decisions affecting them. Yet, AI EdTech can also be used to assist students without analysing them. For instance, automatic captions can help students with disabilities such as hearing impairments,

dyslexia, or ADHD follow and take notes in a lecture [121]. Similarly, ChatGPT can be used to summarise or simplify texts, and thereby facilitate communication and learning for people with disabilities such as dyslexia or speech impairments [122]. While this type of use was out of the scope of this study, readers should bear in mind that AI can be employed as assistive technologies to promote inclusion.

ACKNOWLEDGMENT

The authors would like to thank all the experts who provided them with input on the topic, namely Gerd Kortemeyer, Olivia Heinzer, Benjamin Börner, Daniele Corciulo, Clemens Mader, Ricardo Chavarriga, Corinna Hertweck, and David Mosquera for his support with the method. They would also like to thank Juliet Manning for editing this article.

The authors declare that all ideas and analyses were produced by themselves. However, ChatGPT was employed as a writing assistant tool to improve the writing style of the article. More specifically, some sentences were pasted in ChatGPT with the prompt “reformulate “[sentence]” or “is it correct “[sentence].” The second prompt checked the grammatical correctness of the sentence. The resulting text suggestions were assessed by the Oriane Pierrès, and if they were deemed to be correct or clearer than the original text without modifying its meaning, the suggestion was adopted fully or partially. In addition, ChatGPT was used to help create an interactive dashboard with R. Nevertheless, ChatGPT was not used to provide new ideas nor to inform the content.

REFERENCES

- [1] S. Russell and P. Norvig, *Artificial Intelligence, Global Edition: A Modern Approach*. London, U.K.: Pearson, 2021. [Online]. Available: <https://elibrary.pearson.de/book/99.150005/9781292401171>
- [2] A. Bozkurt, A. Karadeniz, D. Baneres, A. E. Guerrero-Roldán, and M. E. Rodríguez, “Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century,” *Sustainability*, vol. 13, no. 2, p. 800, Jan. 2021.
- [3] V. González-Calatayud, P. Prendes-Espinosa, and R. Roig-Vila, “Artificial intelligence for student assessment: A systematic review,” *Appl. Sci.*, vol. 11, no. 12, p. 5467, Jun. 2021, doi: [10.3390/app11125467](https://doi.org/10.3390/app11125467).
- [4] T. Kabudi, I. Pappas, and D. H. Olsen, “AI-enabled adaptive learning systems: A systematic mapping of the literature,” *Comput. Educ., Artif. Intell.*, vol. 2, Jan. 2021, Art. no. 100017, doi: [10.1016/j.caeai.2021.100017](https://doi.org/10.1016/j.caeai.2021.100017).
- [5] H. Munir, B. Vogel, and A. Jacobsson, “Artificial intelligence and machine learning approaches in digital education: A systematic revision,” *Information*, vol. 13, no. 4, p. 203, Apr. 2022, doi: [10.3390/info13040203](https://doi.org/10.3390/info13040203).
- [6] F. Ouyang, L. Zheng, and P. Jiao, “Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020,” *Educ. Inf. Technol.*, vol. 27, no. 6, pp. 7893–7925, Jul. 2022, doi: [10.1007/s10639-022-10925-9](https://doi.org/10.1007/s10639-022-10925-9).
- [7] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, “Systematic review of research on artificial intelligence applications in higher education—Where are the educators?” *Int. J. Educ. Technol. Higher Educ.*, vol. 16, no. 1, pp. 1–27, Dec. 2019.
- [8] X. Zhai, X. Chu, C. S. Chai, M. S. Y. Jong, and A. Istenic, “A review of artificial intelligence (AI) in education from 2010 to 2020,” *Complexity*, vol. 2021, pp. 1–18, Apr. 2021.
- [9] D. C. Brooks, “Artificial intelligence use in higher education,” in *Educause Review—Special Report. Artificial Intelligence Where are we Now?* Boulder, CO, USA: Educause, 2022, pp. 18–25. [Online]. Available: <https://www.educause.edu/showcase-series/2022/ai-where-are-we-now>

- [10] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *IEEE Access*, vol. 8, pp. 75264–75278, 2020, doi: [10.1109/ACCESS.2020.2988510](https://doi.org/10.1109/ACCESS.2020.2988510).
- [11] Y. Cui, F. Chen, A. Shiri, and Y. Fan, "Predictive analytic models of student success in higher education: A review of methodology," *Inf. Learn. Sci.*, vol. 120, no. 3, pp. 208–227, Mar. 2019, doi: [10.1108/ils-10-2018-0104](https://doi.org/10.1108/ils-10-2018-0104).
- [12] M. Hooda, C. Rana, O. Dahiya, A. Rizwan, and M. S. Hossain, "Artificial intelligence for assessment and feedback to enhance student success in higher education," *Math. Problems Eng.*, vol. 2022, pp. 1–19, May 2022, doi: [10.1155/2022/5215722](https://doi.org/10.1155/2022/5215722).
- [13] E. Okewu, P. Adewole, S. Misra, R. Maskeliunas, and R. Damasevicius, "Artificial neural networks for educational data mining in higher education: A systematic literature review," *Appl. Artif. Intell.*, vol. 35, no. 13, pp. 983–1021, Nov. 2021, doi: [10.1080/08839514.2021.1922847](https://doi.org/10.1080/08839514.2021.1922847).
- [14] A. Nigam, R. Pasricha, T. Singh, and P. Churi, "A systematic review on AI-based proctoring systems: Past, present and future," *Educ. Inf. Technol.*, vol. 26, no. 5, pp. 6421–6445, Sep. 2021, doi: [10.1007/s10639-021-10597-x](https://doi.org/10.1007/s10639-021-10597-x).
- [15] World Health Organisation. *Disability and Health*. Accessed: Feb. 12, 2024. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/disability-and-health#:~:text=Disability%20refers%20to%20the%20interaction,%2C%20and%20limited%20social%20supports>
- [16] CAST. *Universal Design for Learning Guidelines Version 2.2*. Accessed: Feb. 12, 2024. [Online]. Available: <http://udlguidelines.cast.org/>
- [17] J. M. García-González, S. G. Gómez-Calcerrada, E. S. Hernández, and S. Ríos-Aguilar, "Barriers in higher education: Perceptions and discourse analysis of students with disabilities in Spain," *Disability Soc.*, vol. 36, no. 4, pp. 579–595, Apr. 2021, doi: [10.1080/09687599.2020.1749565](https://doi.org/10.1080/09687599.2020.1749565).
- [18] T. Heiman, T. Coughlan, H. Rangin, and M. Deimann, "New designs or new practices? Multiple perspectives on the ICT and accessibility conundrum," in *Improving Accessible Digital Practices in Higher Education: Challenges and New Practices for Inclusion*, J. Seale, Ed. Cham, Switzerland: Springer, 2020, pp. 99–115, doi: [10.1007/978-3-030-37125-8_5](https://doi.org/10.1007/978-3-030-37125-8_5).
- [19] M. Khalil, S. Slade, and P. Prinsloo, "Learning analytics in support of inclusiveness and disabled students: A systematic review," *J. Comput. Higher Educ.*, pp. 1–18, Mar. 2023, doi: [10.1007/s12528-023-09363-4](https://doi.org/10.1007/s12528-023-09363-4).
- [20] M. Hooda, C. Rana, O. Dahiya, J. P. Shet, and B. K. Singh, "Integrating LA and EDM for improving students success in higher education using FCN algorithm," *Math. Problems Eng.*, vol. 2022, pp. 1–12, Mar. 2022, doi: [10.1155/2022/7690103](https://doi.org/10.1155/2022/7690103).
- [21] C. Obeid, C. Lahoud, K. El, and P.-A. Champin, "A novel hybrid recommender system approach for student academic advising named COHRS, supported by case-based reasoning and ontology," *Comput. Sci. Inf. Syst.*, vol. 19, no. 2, pp. 979–1005, 2022, doi: [10.2298/CSIS2202150110](https://doi.org/10.2298/CSIS2202150110).
- [22] B. Heinrichs, "Discrimination in the age of artificial intelligence," *AI Soc.*, vol. 37, no. 1, pp. 143–154, Mar. 2022, doi: [10.1007/s00146-021-01192-2](https://doi.org/10.1007/s00146-021-01192-2).
- [23] M. R. Morris, "AI and accessibility," *Commun. ACM*, vol. 63, no. 6, pp. 35–37, May 2020, doi: [10.1145/3356727](https://doi.org/10.1145/3356727).
- [24] A. Guo, E. Kamar, J. W. Vaughan, H. Wallach, and M. R. Morris, "Toward fairness in AI for people with disabilities SBG@a research roadmap," *ACM SIGACCESS Accessibility Comput.*, no. 125, p. 8, Mar. 2020, Art. no. 2.
- [25] R. S. Baker and A. Hawn, "Algorithmic bias in education," *Int. J. Artif. Intell. Educ.*, vol. 32, no. 4, pp. 1052–1092, Nov. 2021, doi: [10.1007/s40593-021-00285-9](https://doi.org/10.1007/s40593-021-00285-9).
- [26] B. Hutchinson, V. Prabhakaran, E. Denton, K. Webster, Y. Zhong, and S. Denuyl, "Unintended machine learning biases as social barriers for persons with disabilities," *ACM SIGACCESS Accessibility Comput.*, no. 125, p. 10, Mar. 2020, Art. no. 9, doi: [10.1145/3386296.3386305](https://doi.org/10.1145/3386296.3386305).
- [27] S. Trewin, "AI fairness for people with disabilities: Point of view," 2018, *arXiv:1811.10670*.
- [28] H. Profita, R. Albaghli, L. Findlater, P. Jaeger, and S. K. Kane, "The AT effect: How disability affects the perceived social acceptability of head-mounted display use," in *Proc. CHI Conf. Human Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, May 2016, pp. 4884–4895, doi: [10.1145/2858036.2858130](https://doi.org/10.1145/2858036.2858130).
- [29] European Commission. (2019). *Ethics Guidelines for Trustworthy AI Brussels*. [Online]. Available: <https://wayback.archive-it.org/12090/20201227221227/https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- [30] H. Suresh and J. Gutttag, "A framework for understanding sources of harm throughout the machine learning life cycle," in *Equity and Access in Algorithms, Mechanisms, and Optimization*. New York, NY, USA: Association for Computing Machinery (ACM), Oct. 2021, pp. 1–9, doi: [10.1145/3465416.3483305](https://doi.org/10.1145/3465416.3483305).
- [31] M. Brkan, "Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond," *Int. J. Law Inf. Technol.*, vol. 27, no. 2, pp. 91–121, Jun. 2019, doi: [10.1093/ijlit/eay017](https://doi.org/10.1093/ijlit/eay017).
- [32] N. Kordzadeh and M. Ghasemaghaei, "Algorithmic bias: Review, synthesis, and future research directions," *Eur. J. Inf. Syst.*, vol. 31, no. 3, pp. 388–409, Jun. 2021, doi: [10.1080/0960085x.2021.1927212](https://doi.org/10.1080/0960085x.2021.1927212).
- [33] M. J. Page, D. Moher, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamsee, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, and R. Chou, "PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews," *BMJ*, vol. 372, p. 160, Mar. 2021, doi: [10.1136/bmj.n160](https://doi.org/10.1136/bmj.n160).
- [34] A. C. Tricco, E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks, and S. Hempel, "PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation," *Ann. Internal Med.*, vol. 169, no. 7, pp. 467–473, Oct. 2018, doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850).
- [35] A. Moriña, "Inclusive education in higher education: Challenges and opportunities," *Eur. J. Special Needs Educ.*, vol. 32, no. 1, pp. 3–17, Jan. 2017, doi: [10.1080/08856257.2016.1254964](https://doi.org/10.1080/08856257.2016.1254964).
- [36] T. Robal, Y. Zhao, C. Lofi, and C. Hauff, "IntelliEye: Enhancing MOOC learners' video watching experience through real-time attention tracking," in *Proc. 29th Hypertext Social Media*, Baltimore MD USA, 2018, pp. 106–114, doi: [10.1145/3209542.3209547](https://doi.org/10.1145/3209542.3209547).
- [37] Q. Jia, M. Young, Y. Xiao, J. Cui, C. Liu, P. Rashid, and E. Gehringer, "Insta-reviewer: A data-driven approach for generating instant feedback on students' project reports," in *Proc. 15th Int. Conf. Educ. Data Mining*, Durham, U.K., 2022, pp. 5–16.
- [38] X. Wei, S. Sun, D. Wu, and L. Zhou, "Personalized online learning resource recommendation based on artificial intelligence and educational psychology," *Frontiers Psychol.*, vol. 12, Dec. 2021, Art. no. 767837, doi: [10.3389/fpsyg.2021.767837](https://doi.org/10.3389/fpsyg.2021.767837).
- [39] P. Hur, H. Lee, S. Bhat, and N. Bosch, "Using machine learning explainability methods to personalize interventions for students," in *Proc. 15th Int. Conf. Educ. Data Mining*, Durham, U.K., 2022, pp. 438–445.
- [40] M. Nagy and R. Molontay, "Interpretable dropout prediction: Towards XAI-based personalized intervention," *Int. J. Artif. Intell. Educ.*, pp. 1–2, Mar. 2023, doi: [10.1007/s40593-023-00331-8](https://doi.org/10.1007/s40593-023-00331-8).
- [41] J. W. Orr and N. Russell, "Automatic assessment of the design quality of python programs with personalized feedback," in *Proc. 14th Int. Conf. Educ. Data Mining (EDM)*, Paris, France, 2021, pp. 495–501. [Online]. Available: <https://educationaldatamining.org/edm2021/>
- [42] A. Ortigosa, R. M. Carro, J. Bravo-Agapito, D. Lizcano, J. J. Alcolea, and Ó. Blanco, "From lab to production: Lessons learnt and real-life challenges of an early student-dropout prevention system," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 264–277, Apr. 2019, doi: [10.1109/TLT.2019.2911608](https://doi.org/10.1109/TLT.2019.2911608).
- [43] L. Roux, T. Nodenot, P. Etcheverry, P. Dagorret, C. Marquesuzaa, and P. Lopisteguy, "A learner's behavior model for an e-learning hybrid recommender system," in *Open and Inclusive Educational Practice in the Digital World (Cognition and Exploratory Learning in the Digital Age)*, D. Ifenthaler, D. G. Sampson, and P. Isaías, Eds. Cham, Switzerland: Springer, 2022, pp. 65–81, doi: [10.1007/978-3-031-18512-0_5](https://doi.org/10.1007/978-3-031-18512-0_5).
- [44] D. Azcona, I.-H. Hsiao, P. Arora, and A. Smeaton, "User2Code2vec: Embeddings for profiling students based on distributional representations of source code," in *Proc. ACM Int. Conf.*, 2019, pp. 86–95, doi: [10.1145/3303772.3303813](https://doi.org/10.1145/3303772.3303813).
- [45] L. Aulck, D. Nambi, and J. West, "Increasing enrollment by optimizing scholarship allocations using machine learning and genetic algorithms," in *Proc. 13th Int. Conf. Educ. Data Mining (EDM)*, 2020, pp. 29–38.
- [46] A.-E. Guerrero-Roldán, M. E. Rodríguez-González, D. Bañeres, A. Elasmri-Ejjaberi, and P. Cortadas, "Experiences in the use of an adaptive intelligent system to enhance online learners' performance: A case study in economics and business courses," *Int. J. Educ. Technol. Higher Educ.*, vol. 18, no. 1, pp. 1–27, Jul. 2021, doi: [10.1186/s41239-021-00271-0](https://doi.org/10.1186/s41239-021-00271-0).
- [47] E. Demeter, M. Dorodchi, E. Al-Hossami, A. Benedict, L. Slattery Walker, and J. Smail, "Predicting first-time-in-college students' degree completion outcomes," *Higher Educ.*, vol. 84, no. 3, pp. 589–609, Sep. 2022, doi: [10.1007/s10734-021-00790-9](https://doi.org/10.1007/s10734-021-00790-9).

- [48] J. Renzella, A. Cain, and J. Schneider, "An intelligent tool for combating contract cheating behaviour by facilitating scalable student-tutor discussions," in *Proc. ACM/IEEE 42nd Int. Conf. Softw. Eng.*, Jun. 2020, pp. 298–299, doi: [10.1145/3377812.3390795](https://doi.org/10.1145/3377812.3390795).
- [49] S.-C. Tsai, C.-H. Chen, Y.-T. Shiao, J.-S. Ciou, and T.-N. Wu, "Precision education with statistical learning and deep learning: A case study in Taiwan," *Int. J. Educ. Technol. Higher Educ.*, vol. 17, no. 1, pp. 1–13, Apr. 2020, doi: [10.1186/s41239-020-00186-2](https://doi.org/10.1186/s41239-020-00186-2).
- [50] X. Xia, "Interaction recognition and intervention based on context feature fusion of learning behaviors in interactive learning environments," *Interact. Learn. Environ.*, vol. 31, no. 4, pp. 2033–2050, Jan. 2021, doi: [10.1080/10494820.2021.1871632](https://doi.org/10.1080/10494820.2021.1871632).
- [51] L. Z. Aljuaid and M. Y. Said, "Deep learning-based method for prediction of software engineering project teamwork assessment in higher education," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 9, pp. 2012–2030, 2021.
- [52] M. Ciolacu, A. F. Tehrani, L. Binder, and P. M. Svasta, "Education 4.0—Artificial intelligence assisted higher education: Early recognition system with machine learning to support students' success," in *Proc. IEEE 24th Int. Symp. Design Technol. Electron. Packag. (SIITME)*, Oct. 2018, pp. 23–30.
- [53] E. de Quincey, C. Briggs, T. Kyriacou, and R. Waller, "Student centred design of a learning analytics system," in *Proc. 9th Int. Conf. Learn. Analytics Knowl.* New York, NY, USA: Association for Computing Machinery, Mar. 2019, pp. 353–362, doi: [10.1145/3303772.3303793](https://doi.org/10.1145/3303772.3303793).
- [54] S. B. Dias, S. J. Hadjileontiadou, J. Diniz, and L. J. Hadjileontiadis, "DeepLMS: A deep learning predictive model for supporting online learning in the COVID-19 era," *Sci. Rep.*, vol. 10, no. 1, p. 19888, Nov. 2020, doi: [10.1038/s41598-020-76740-9](https://doi.org/10.1038/s41598-020-76740-9).
- [55] M. Dorodchi, A. Benedict, D. Desai, M. J. Mahzoon, S. MacNeil, and N. Dehbozorgi, "Design and implementation of an activity-based introductory computer science course (CS1) with periodic reflections validated by learning analytics," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, San Jose, CA, USA, Oct. 2018, pp. 1–8, doi: [10.1109/FIE.2018.8659196](https://doi.org/10.1109/FIE.2018.8659196).
- [56] M. Eagle, T. Carmichael, J. Stokes, M. J. Blink, J. C. Stamper, and J. Levin, "Predictive student modeling for interventions in online classes," in *Proc. 11th Int. Conf. Educ. Data Mining*, K. E. Boyer and M. Yudelson, Eds. Buffalo, NY, USA, Jul. 2018, pp. 619–624. [Online]. Available: <https://api.semanticscholar.org/CorpusID:52175577>
- [57] J. Figueiredo and F. García-Peñalvo, "A tool help for introductory programming courses," in *Proc. ACM Int. Conf.*, 2021, pp. 18–24, doi: [10.1145/3486011.3486413](https://doi.org/10.1145/3486011.3486413).
- [58] J. Hellings and C. Haelermans, "The effect of providing learning analytics on student behaviour and performance in programming: A randomised controlled experiment," *Higher Educ.*, vol. 83, no. 1, pp. 1–18, Jan. 2022, doi: [10.1007/s10734-020-00560-z](https://doi.org/10.1007/s10734-020-00560-z).
- [59] M. Hussain, W. Zhu, W. Zhang, and S. M. R. Abidi, "Student engagement predictions in an e-learning system and their impact on student course assessment scores," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–21, Oct. 2018, doi: [10.1155/2018/6347186](https://doi.org/10.1155/2018/6347186).
- [60] C. V. Le, Z. A. Pardos, S. D. Meyer, and R. Thorp, "Communication at scale in a MOOC using predictive engagement analytics," in *Artificial Intelligence in Education* (Lecture Notes in Computer Science), vol. 10947, C. P. Rose, R. Martinez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, and B. du Boulay, Eds. Cham, Switzerland: Springer, 2018, pp. 239–252, doi: [10.1007/978-3-319-93843-1_18](https://doi.org/10.1007/978-3-319-93843-1_18).
- [61] D. Monllaó Olivé, D. Q. Huynh, M. Reynolds, M. Dougiamas, and D. Wiese, "A supervised learning framework: Using assessment to identify students at risk of dropping out of a MOOC," *J. Comput. Higher Educ.*, vol. 32, no. 1, pp. 9–26, Apr. 2020, doi: [10.1007/s12528-019-09230-1](https://doi.org/10.1007/s12528-019-09230-1).
- [62] C. Van Petegem, L. Deconinck, D. Mourisse, R. Maertens, N. Strijbol, B. Dhoedt, B. De Wever, P. Dawyndt, and B. Mesuere, "Pass/fail prediction in programming courses," *J. Educ. Comput. Res.*, vol. 61, no. 1, pp. 68–95, Mar. 2023, doi: [10.1177/07356331221085595](https://doi.org/10.1177/07356331221085595).
- [63] W. Xing and D. Du, "Dropout prediction in MOOCs: Using deep learning for personalized intervention," *J. Educ. Comput. Res.*, vol. 57, no. 3, pp. 547–570, Jun. 2019, doi: [10.1177/0735633118757015](https://doi.org/10.1177/0735633118757015).
- [64] A. Baskota and Y.-K. Ng, "A graduate school recommendation system using the multi-class support vector machine and KNN approaches," in *Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI)*, Jul. 2018, pp. 277–284, doi: [10.1109/IRI.2018.00050](https://doi.org/10.1109/IRI.2018.00050).
- [65] S. S. Bhaskaran and M. A. Aali, "Data mining model for better admissions in higher educational institutions (HEIs)—A case study of Bahrain," in *Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl.*, vol. 1141, 2021, p. 150, doi: [10.1007/978-981-15-3383-9_13](https://doi.org/10.1007/978-981-15-3383-9_13).
- [66] B. M. da Silva and C. M. Antunes, "ESTHER: A recommendation system for higher education programs," in *Proc. EPIA Conf. Artif. Intell.*, 2021, pp. 537–548, doi: [10.1007/978-3-030-86230-5_42](https://doi.org/10.1007/978-3-030-86230-5_42).
- [67] A. Nalawade and B. Tiple, "University recommendation system for higher studies in USA," in *ICDSMLA* (Lecture Notes in Electrical Engineering), vol. 783. Singapore: Springer, 2022, pp. 145–153, doi: [10.1007/978-981-16-3690-5_14](https://doi.org/10.1007/978-981-16-3690-5_14).
- [68] O. Akhrif, C. Benfaress, M. El Jai, Y. E. B. El Idrissi, and N. Hmina, "Completeness based classification algorithm: A novel approach for educational semantic data completeness assessment," *Interact. Technol. Smart Educ.*, vol. 19, no. 1, pp. 87–111, Feb. 2022, doi: [10.1108/itse-01-2021-0017](https://doi.org/10.1108/itse-01-2021-0017).
- [69] R. Haddad and E. Mercier-Laurent, "Curriculum vitae (CVs) evaluation using machine learning approach," in *Proc. IFIP Int. Workshop Artif. Intell. Knowl. Manag.*, E. Mercier-Laurent, M. Kayalica, and M. Owoc, Eds. 2021, pp. 48–65, doi: [10.1007/978-3-030-80847-1_4](https://doi.org/10.1007/978-3-030-80847-1_4).
- [70] H. A. Mengash, "Using data mining techniques to predict student performance to support decision making in university admission systems," *IEEE Access*, vol. 8, pp. 55462–55470, 2020, doi: [10.1109/ACCESS.2020.2981905](https://doi.org/10.1109/ACCESS.2020.2981905).
- [71] R. Ajoodha, S. Dukhan, and A. Jadhav, "Data-driven student support for academic success by developing student skill profiles," in *Proc. 2nd Int. Multidisciplinary Inf. Technol. Eng. Conf. (IMITEC)*, Nov. 2020, pp. 1–8, doi: [10.1109/IMITEC50163.2020.9334109](https://doi.org/10.1109/IMITEC50163.2020.9334109).
- [72] S. Yanta, S. Thammaboosadee, P. Chanyagorn, and R. Chuckpaiwong, "Probation status prediction and optimization for undergraduate engineering students," in *Proc. 13th Int. Conf. Knowl. Smart Technol. (KST)*, Jan. 2021, pp. 191–196, doi: [10.1109/KST51265.2021.9415762](https://doi.org/10.1109/KST51265.2021.9415762).
- [73] J. Andrews-Todd, J. Steinberg, M. Flor, and C. M. Forsyth, "Exploring automated classification approaches to advance the assessment of collaborative problem solving skills," *J. Intell.*, vol. 10, no. 3, p. 39, Jul. 2022, doi: [10.3390/jintelligence10030039](https://doi.org/10.3390/jintelligence10030039).
- [74] P. M. Davies, R. J. Passonneau, S. Muresan, and Y. Gao, "Analytical techniques for developing argumentative writing in STEM: A pilot study," *IEEE Trans. Educ.*, vol. 65, no. 3, pp. 373–383, Aug. 2022, doi: [10.1109/TE.2021.3116202](https://doi.org/10.1109/TE.2021.3116202).
- [75] M. Holmes, A. Latham, K. Crockett, and J. D. O'Shea, "Near real-time comprehension classification with artificial neural networks: Decoding e-learner non-verbal behavior," *IEEE Trans. Learn. Technol.*, vol. 11, no. 1, pp. 5–12, Jan. 2018, doi: [10.1109/TLT.2017.2754497](https://doi.org/10.1109/TLT.2017.2754497).
- [76] M. Jahnke and F. Höppner, "Is there method in your mistakes? Capturing error contexts by graph mining for targeted feedback," in *Proc. 15th Int. Conf. Educ. Data Mining*, Durham, U.K., 2022, pp. 422–429.
- [77] A. Malik, M. Wu, V. Vasavada, J. Song, M. Coots, J. Mitchell, N. Goodman, and C. Piech, "Generative grading: Near human-level accuracy for automated feedback on richly structured problems," in *Proc. 14th Int. Conf. Educ. Data Mining (EDM)*, 2021, pp. 1–12.
- [78] D. Miao, Y. Dong, and X. Lu, "PIPE: Predicting logical programming errors in programming exercises," in *Proc. 13th Int. Conf. Educ. Data Mining*, 2020, pp. 473–479.
- [79] A. Poce, C. de Medío, F. Amenduni, and M. R. Re, "Critical thinking assessment: A first approach to the automatic evaluation," in *Proc. 18th Int. Conf. Inf. Technol. Based Higher Educ. Training (ITHET)*, Sep. 2019, pp. 1–8.
- [80] J. Rodríguez-Ruiz, A. Alvarez-Delgado, and P. Caratozzolo, "Use of natural language processing (NLP) tools to assess digital literacy skills," in *Proc. Mach. Learn.-Driven Digital Technol. Educ. Innov. Workshop*, 2021, pp. 1–8, doi: [10.1109/IEEECONF53024.2021.9733779](https://doi.org/10.1109/IEEECONF53024.2021.9733779).
- [81] B. T. Shobana and G. A. S. Kumar, "I-quiz: An intelligent assessment tool for non-verbal behaviour detection," *Comput. Syst. Sci. Eng.*, vol. 40, no. 3, pp. 1007–1021, 2022, doi: [10.32604/csse.2022.019523](https://doi.org/10.32604/csse.2022.019523).
- [82] L. Wang and S. Chen, "Student physical fitness test system and test data analysis system based on computer vision," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–8, May 2021, doi: [10.1155/2021/5589065](https://doi.org/10.1155/2021/5589065).
- [83] B. N. Anh, N. T. Son, P. T. Lam, L. P. Chi, N. H. Tuan, N. C. Dat, N. H. Trung, M. U. Aftab, and T. Van Dinh, "A computer-vision based application for student behavior monitoring in classroom," *Appl. Sci.*, vol. 9, no. 22, p. 4729, Nov. 2019, doi: [10.3390/app9224729](https://doi.org/10.3390/app9224729).

- [84] I. K. Seneviratne, B. A. S. D. Perera, R. S. D. Fernando, L. K. B. Siriwardana, and U. U. S. K. Rajapaksha, "Student and lecturer performance enhancement system using artificial intelligence," in *Proc. 3rd Int. Conf. Intell. Sustain. Syst. (ICISS)*, Dec. 2020, pp. 88–93, doi: [10.1109/ICISS49785.2020.9315981](https://doi.org/10.1109/ICISS49785.2020.9315981).
- [85] J. Wang and Z. Yu, "Smart educational learning strategy with the Internet of Things in higher education system," *Int. J. Artif. Intell. Tools*, vol. 31, Aug. 2021, Art. no. 2140101, doi: [10.1142/S0218213021401011](https://doi.org/10.1142/S0218213021401011).
- [86] F. Kamalov, H. Sulieman, and D. Santandreu Calonge, "Machine learning based approach to exam cheating detection," *PLoS ONE*, vol. 16, no. 8, Aug. 2021, Art. no. e0254340.
- [87] L. K. Smirani and J. A. Boulahia, "An algorithm based on convolutional neural networks to manage online exams via learning management system without using a webcam," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 3, pp. 290–299, 2022, doi: [10.14569/ijacsa.2022.0130336](https://doi.org/10.14569/ijacsa.2022.0130336).
- [88] D. Sunaryono, J. Siswanto, and R. Anggoro, "An Android based course attendance system using face recognition," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 33, no. 3, pp. 304–312, Mar. 2021, doi: [10.1016/j.jksuci.2019.01.006](https://doi.org/10.1016/j.jksuci.2019.01.006).
- [89] Y. Zhou and L. Li, "Research on recommendation of college mental health teaching materials based on improved deep learning algorithm," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–11, Jun. 2022, doi: [10.1155/2022/5184221](https://doi.org/10.1155/2022/5184221).
- [90] Y. Zhu, "The impact of digital technology on the optimization of higher education teaching models in an epidemic environment," *Secur. Commun. Netw.*, vol. 2022, pp. 1–9, May 2022, doi: [10.1155/2022/5301459](https://doi.org/10.1155/2022/5301459).
- [91] K. Dahdouh, A. Dakkak, L. Oughdir, and A. Ibriz, "Large-scale e-learning recommender system based on Spark and Hadoop," *J. Big Data*, vol. 6, no. 1, pp. 1–23, Dec. 2019, doi: [10.1186/s40537-019-0169-4](https://doi.org/10.1186/s40537-019-0169-4).
- [92] R. Morsomme and S. V. Alferex, "Content-based course recommender system for Liberal Arts education," in *Proc. 12th Int. Conf. Educ. Data Mining (EDM)*, C. F. Lynch, A. Merceron, M. Desmarais, and R. Nkam-bou, Eds., Jul. 2019, pp. 748–753.
- [93] Z. Chen and C. Demmans, "CSCLRec: Personalized recommendation of forum posts to support socio-collaborative Learning," in *Proc. The 13th Int. Conf. Educ. Data Mining*, 2020, pp. 364–373.
- [94] H. Chen, C. Yin, X. Fan, L. Qiao, W. Rong, and X. Zhang, "Learning path recommendation for MOOC platforms based on a knowledge graph," in *Proc. Int. Conf. Knowl. Sci., Eng. Manag.*, H. Qiu, C. Zhang, Z. Fei, M. Qiu, and S. Kung, Eds. 2021, pp. 600–611, doi: [10.1007/978-3-030-82147-0_49](https://doi.org/10.1007/978-3-030-82147-0_49).
- [95] D. Datta, M. Phillips, J. P. Bywater, J. Chiu, G. S. Watson, L. Barnes, and D. Brown, "Virtual pre-service teacher assessment and feedback via conversational agents," in *Proc. 16th Workshop Innov. Use NLP Building Educ. Appl.*, 2021, pp. 185–198. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85109544516&partnerID=40&md5=8f8e63b1cf0a2d95732ebe57cef2740c>
- [96] D. Rossi, V. Stroele, R. Braga, S. Caballe, N. Capuano, F. Campos, M. Dantas, L. Lomasto, and D. Toti, "CAERS: A conversational agent for intervention in MOOCs' learning processes," in *Innovations in Learning and Technology for the Workplace and Higher Education* (Lecture Notes in Networks and Systems), vol. 349, D. Guralnick, M. Auer, and A. Poce, Eds. Cham, Switzerland: Springer, 2022, pp. 371–382, doi: [10.1007/978-3-030-90677-1_36](https://doi.org/10.1007/978-3-030-90677-1_36).
- [97] T. Schlippe and J. Sawatzki, "AI-based multilingual interactive exam preparation," in *Innovations in Learning and Technology for the Workplace and Higher Education* (Lecture Notes in Networks and Systems), vol. 349, D. Guralnick, M. Auer, and A. Poce, Eds. Cham, Switzerland: Springer, 2022, pp. 396–408, doi: [10.1007/978-3-030-90677-1_38](https://doi.org/10.1007/978-3-030-90677-1_38).
- [98] H. C. R. Jeewantha, A. N. Gajasinghe, N. I. Naidabadu, T. N. Rajapaksha, D. Kasthurirathna, and A. Karunasena, "English language trainer for non-native speakers using audio signal processing, reinforcement learning, and deep learning," in *Proc. 21st Int. Conf. Adv. ICT Emerg. Regions (ICter)*, Dec. 2021, pp. 117–122, doi: [10.1109/ICter53630.2021.9774785](https://doi.org/10.1109/ICter53630.2021.9774785).
- [99] M. Kim, N. Kim, and A. Heidari, "Learner experience in artificial intelligence-scaffolded argumentation," *Assessment Eval. Higher Educ.*, vol. 47, no. 8, pp. 1301–1316, 2022, doi: [10.1080/02602938.2022.2042792](https://doi.org/10.1080/02602938.2022.2042792).
- [100] T. W. Li and L. Paquette, "Erroneous answers categorization for sketching questions in spatial visualization training," in *Proc. 13th Int. Conf. Educ. Data Mining*, 2020, p. 5068.
- [101] O. Lyashevskaya, I. Panteleeva, and O. Vinogradova, "Automated assessment of learner text complexity," *Assessing Writing*, vol. 49, Jul. 2021, Art. no. 100529, doi: [10.1016/j.asw.2021.100529](https://doi.org/10.1016/j.asw.2021.100529).
- [102] S. Marwan, Y. Shi, I. Menezes, M. Chi, T. Barnes, and T. W. Price, "Just a few expert constraints can help: Humanizing data-driven subgoal detection for novice programming," in *Proc. 14th Int. Conf. Educ. Data Mining*, Paris, France, 2021, pp. 68–80.
- [103] M. Sanz Ausin, M. Maniktala, T. Barnes, and M. Chi, "Exploring the impact of simple explanations and agency on batch deep reinforcement learning induced pedagogical policies," in *Artificial Intelligence in Education*, I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, and E. Millán, Eds. Cham, Switzerland: Springer, 2020, pp. 472–485.
- [104] A. Unnam, R. Takhar, and V. Aggarwal, "Grading emails and generating feedback," in *Proc. 12th Int. Conf. Educ. Data Mining*, 2019, pp. 1–6.
- [105] J. M. Vytasek, A. Patzak, and P. H. Winne, "Topic development to support revision feedback," in *Proc. 9th Int. Conf. Learn. Analytics Knowl.* New York, NY, USA: Association for Computing Machinery, 2019, pp. 220–224, doi: [10.1145/3303772.3303816](https://doi.org/10.1145/3303772.3303816).
- [106] C. Baek and S. J. Aguilar, "Past, present, and future directions of learning analytics research for students with disabilities," *J. Res. Technol. Educ.*, vol. 55, no. 6, pp. 931–946, May 2022, doi: [10.1080/15391523.2022.2067796](https://doi.org/10.1080/15391523.2022.2067796).
- [107] W. Chen, "Learning analytics for inclusive higher education," in *Proc. 28th Int. Conf. Comput. Educ.*, 2020, pp. 113–118.
- [108] S. Riazzy, K. Simbeck, and V. Schreck, "Fairness in learning analytics: Student at-risk prediction in virtual learning environments," in *Proc. 12th Int. Conf. Comput. Supported Educ.*, 2020, pp. 15–25, doi: [10.5220/0009324100150025](https://doi.org/10.5220/0009324100150025).
- [109] M. Cooper, R. Ferguson, and A. Wolff, "What can analytics contribute to accessibility in e-learning systems and to disabled students' learning?" in *Proc. 6th Int. Conf. Learn. Analytics Knowl.*, 2016, pp. 99–103, doi: [10.1145/2883851.2883946](https://doi.org/10.1145/2883851.2883946).
- [110] N. Selwyn, "Re-imagining 'learning analytics'...a case for starting again?" *Internet Higher Educ.*, vol. 46, Jul. 2020, Art. no. 100745, doi: [10.1016/j.iheduc.2020.100745](https://doi.org/10.1016/j.iheduc.2020.100745).
- [111] F. Steinhart, A. Ullenhag, R. Jahnsen, and A.-S. Dolva, "Perceived facilitators and barriers for participation in leisure activities in children with disabilities: Perspectives of children, parents and professionals," *Scandin. J. Occupational Therapy*, vol. 28, no. 2, pp. 121–135, Feb. 2021, doi: [10.1080/11038128.2019.1703037](https://doi.org/10.1080/11038128.2019.1703037).
- [112] T. R. Wolanin, "Students with disabilities: Financial aid policy issues," *J. Student Financial Aid*, vol. 35, no. 1, p. 2, Feb. 2005.
- [113] K. M. Soria, B. Horgos, and J. D. Shenouda, "Disparities in college students' financial hardships during the COVID-19 pandemic," *J. Student Affairs Res. Pract.*, vol. 60, no. 1, pp. 31–48, Jan. 2023, doi: [10.1080/19496591.2022.2046597](https://doi.org/10.1080/19496591.2022.2046597).
- [114] D. Chiwandire and L. Vincent, "Funding and inclusion in higher education institutions for students with disabilities," *Afr. J. Disability*, vol. 8, pp. 1–12, Jan. 2019, doi: [10.4102/ajod.v8i0.336](https://doi.org/10.4102/ajod.v8i0.336).
- [115] L. Mullins and M. Preyde, "The lived experience of students with an invisible disability at a Canadian University," *Disability Soc.*, vol. 28, no. 2, pp. 147–160, Mar. 2013, doi: [10.1080/09687599.2012.752127](https://doi.org/10.1080/09687599.2012.752127).
- [116] T. Osborne, "Not lazy, not faking: Teaching and learning experiences of university students with disabilities," *Disability Soc.*, vol. 34, no. 2, pp. 228–252, Feb. 2019, doi: [10.1080/09687599.2018.1515724](https://doi.org/10.1080/09687599.2018.1515724).
- [117] I. D. Raji, I. E. Kumar, A. Horowitz, and A. Selbst, "The fallacy of AI functionality," in *Proc. ACM Conf. Fairness, Accountability Transparency*, Jun. 2022, pp. 959–972, doi: [10.1145/3531146.3533158](https://doi.org/10.1145/3531146.3533158).
- [118] S. Strauß, "Deep automation bias: How to tackle a wicked problem of AI?" *Big Data Cognit. Comput.*, vol. 5, no. 2, p. 18, Apr. 2021, doi: [10.3390/bdcc5020018](https://doi.org/10.3390/bdcc5020018).
- [119] (2004). *Individuals With Disabilities Act (IDEA)*. [Online]. Available: <https://sites.ed.gov/idea/statute-chapter-33/subchapter-i/1401/1>
- [120] E. Morozov. *Avoiding Solutionism in the Digital Transformation of Education*. UNESCO Futures Educ. Ideas LAB. [Online]. Available: <https://en.unesco.org/futuresofeducation/ideas-lab/morozov-avoid-solutionism-digital-transformation>
- [121] M. Kent, K. Ellis, N. Latter, and G. Peaty, "The case for captioned lectures in Australian higher education," *TechTrends*, vol. 62, no. 2, pp. 158–165, Mar. 2018, doi: [10.1007/s11528-017-0225-x](https://doi.org/10.1007/s11528-017-0225-x).
- [122] B. Hemsley, E. Power, and F. Given. *Will AI Tech Like ChatGPT Improve Inclusion for People With Communication Disability?* Accessed: Mar. 21, 2023. [Online]. Available: <https://theconversation.com/will-ai-tech-like-chatgpt-improve-inclusion-for-people-with-communication-disability-196481>

- [123] O. Pierrès, A. Darvishy, and M. Christen, “Artificial intelligence in higher education: Ethical concerns for students with disabilities,” in *Proc. 2nd European Workshop Algorithmic Fairness*, Winterthur, Switzerland, Jun. 2023, pp. 1–5. [Online]. Available: <https://ceur-ws.org/Vol-3442/paper-37.pdf>



ORIANE PIERRÈS received the bachelor’s degree in European studies from Clermont Auvergne University, France, and the M.A. degree in politics and public administration from the University of Konstanz, Germany. She is currently pursuing the Ph.D. degree in biomedical ethics and law with the Institute of Biomedical Ethics (IBME), University of Zurich (UZH), Switzerland. She is a Research Assistant with the ICT-Accessibility Laboratory, Zurich University of Applied Sciences (ZHAW), Winterthur. Her research interest includes digital accessibility, the use of AI in higher education and students with disabilities.



MARKUS CHRISTEN is currently the Research Group Leader with the IBME, University of Zurich, and the Managing Director of the UZH Digital Society Initiative. His research interests include empirical ethics, neuroethics, ethics and technology, serious moral games, and data analysis methodologies.



FELIX M. SCHMITT-KOOPMANN received the B.Sc. degree in mechanical engineering and the M.Sc. degree in robotics, systems, and control from ETH Zürich, in 2019 and 2021, respectively. He is currently pursuing the Ph.D. degree with the People and Computing Laboratory, University of Zurich. He is a member of the Institute of Applied Informatics, Zurich University of Applied Sciences (ZHAW). His research interests include accessibility, AI, and document analysis.



ALIREZA DARVISHY is currently a Professor and the Head of the ICT Accessibility Laboratory, Zurich University of Applied Sciences (ZHAW), Switzerland. He serves as an independent reviewer for European research projects, such as the Active Assisted Living (AAL) Program. He is a principal investigator of the “Accessible Teaching in Higher Education” Project and the “Mastering Accessibility of Open Educational Resources,” funded by swissuniversities.

...