Systematic review of research on artificial intelligence applications in higher education – where are the educators?

Olaf Zawacki-Richter[*] , Victoria I. Marín , Melissa Bond and Franziska Gouverneur

* Correspondence: olaf.zawacki.
richter@uni-oldenburg.de
Faculty of Education and Social
Sciences, University of Oldenburg,
Ammerländer Heerstr. 138, 26129
Oldenburg, Germany

Abstract

According to various international reports, Artificial Intelligence in Education (AIEd) is one of the currently emerging fields in educational technology. Whilst it has been around for about 30 years, it is still unclear for educators how to make pedagogical advantage of it on a broader scale, and how it can actually impact meaningfully on teaching and learning in higher education. This paper seeks to provide an overview of research on AI applications in higher education through a systematic review. Out of 2656 initially identified publications for the period between 2007 and 2018, 146 articles were included for final synthesis, according to explicit inclusion and exclusion criteria. The descriptive results show that most of the disciplines involved in AIEd papers come from Computer Science and STEM, and that quantitative methods were the most frequently used in empirical studies. The synthesis of results presents four areas of AIEd applications in academic support services, and institutional and administrative services: 1. profiling and prediction, 2. assessment and evaluation, 3. adaptive systems and personalisation, and 4. intelligent tutoring systems. The conclusions reflect on the almost lack of critical reflection of challenges and risks of AIEd, the weak connection to theoretical pedagogical perspectives, and the need for further exploration of ethical and educational approaches in the application of AIEd in higher education.

Keywords: Artificial intelligence, Higher education, Machine learning, Intelligent tutoring systems, Systematic review

Introduction

Artificial intelligence (AI) applications in education are on the rise and have received a lot of attention in the last couple of years. AI and adaptive learning technologies are prominently featured as important developments in educational technology in the 2018 Horizon report (Educause, 2018), with a time to adoption of 2 or 3 years. According to the report, experts anticipate AI in education to grow by 43% in the period 2018–2022, although the Horizon Report 2019 Higher Education Edition (Educause, 2019) predicts that AI applications related to teaching and learning are projected to grow even more significantly than this. Contact North, a major Canadian non-profit online learning society, concludes that "there is little doubt that the [AI] technology is inexorably linked to the future of higher education" (Contact North, 2018, p. 5). With heavy investments by private companies such as Google, which acquired European AI

start-up Deep Mind for $400 million, and also non-profit public-private partnerships such as the German Research Centre for Artificial Intelligence[1] (DFKI), it is very likely that this wave of interest will soon have a significant impact on higher education institutions (Popenici & Kerr, 2017). The Technical University of Eindhoven in the Netherlands, for example, recently announced that they will launch an Artificial Intelligence Systems Institute with 50 new professorships for education and research in AI.[2]

The application of AI in education (AIEd) has been the subject of research for about 30 years. The International AIEd Society (IAIED) was launched in 1997, and publishes the International Journal of AI in Education (IJAIED), with the 20th annual AIEd conference being organised this year. However, on a broader scale, educators have just started to explore the potential pedagogical opportunities that AI applications afford for supporting learners during the student life cycle.

Despite the enormous opportunities that AI might afford to support teaching and learning, new ethical implications and risks come in with the development of AI applications in higher education. For example, in times of budget cuts, it might be tempting for administrators to replace teaching by profitable automated AI solutions. Faculty members, teaching assistants, student counsellors, and administrative staff may fear that intelligent tutors, expert systems and chat bots will take their jobs. AI has the potential to advance the capabilities of learning analytics, but on the other hand, such systems require huge amounts of data, including confidential information about students and faculty, which raises serious issues of privacy and data protection. Some institutions have recently been established, such as the Institute for Ethical AI in Education[3] in the UK, to produce a framework for ethical governance for AI in education, and the Analysis & Policy Observatory published a discussion paper in April 2019 to develop an AI ethics framework for Australia.[4]

Russel and Norvig (2010) remind us in their leading textbook on artificial intelligence, "All AI researchers should be concerned with the ethical implications of their work" (p. 1020). Thus, we would like to explore what kind of fresh ethical implications and risks are reflected by the authors in the field of AI enhanced education. The aim of this article is to provide an overview for educators of research on AI applications in higher education. Given the dynamic development in recent years, and the growing interest of educators in this field, a review of the literature on AI in higher education is warranted.

Specifically, this paper addresses the following research questions in three areas, by means of a systematic review (see Gough, Oliver, & Thomas, 2017; Petticrew & Roberts, 2006):

- How have publications on AI in higher education developed over time, in which journals are they published, and where are they coming from in terms of geographical distribution and the author's disciplinary affiliations?

---

[1]https://www.dfki.de/en/web/ (accessed 22 July, 2019)
[2]https://www.tue.nl/en/news/news-overview/11-07-2019-tue-announces-eaisi-new-institute-for-intelligent-machines/ (accessed 22 July, 2019)
[3]http://instituteforethicalaiineducation.org (accessed 22 July, 2019)
[4]https://apo.org.au/node/229596 (accessed 22 July, 2019)

- How is AI in education conceptualised and what kind of ethical implications, challenges and risks are considered?
- What is the nature and scope of AI applications in the context of higher education?

The field AI originates from computer science and engineering, but it is strongly influenced by other disciplines such as philosophy, cognitive science, neuroscience, and economics. Given the interdisciplinary nature of the field, there is little agreement among AI researchers on a common definition and understanding of AI – and intelligence in general (see Tegmark, 2018). With regard to the introduction of AI-based tools and services in higher education, Hinojo-Lucena, Aznar-Díaz, Cáceres-Reche, and Romero-Rodríguez (2019) note that "this technology [AI] is already being introduced in the field of higher education, although many teachers are unaware of its scope and, above all, of what it consists of" (p. 1). For the purpose of our analysis of artificial intelligence in higher education, it is desirable to clarify terminology. Thus, in the next section, we explore definitions of AI in education, and the elements and methods that AI applications might entail in higher education, before we proceed with the systematic review of the literature.

AI in education (AIEd)

The birth of AI goes back to the 1950s when John McCarthy organised a two-month workshop at Dartmouth College in the USA. In the workshop proposal, McCarthy used the term artificial intelligence for the first time in 1956 (Russel & Norvig, 2010, p. 17):

The study [of artificial intelligence] is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.

Baker and Smith (2019) provide a broad definition of AI: "Computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem-solving" (p. 10). They explain that AI does not describe a single technology. It is an umbrella term to describe a range of technologies and methods, such as machine learning, natural language processing, data mining, neural networks or an algorithm.

AI and machine learning are often mentioned in the same breath. Machine learning is a method of AI for supervised and unsupervised classification and profiling, for example to predict the likelihood of a student to drop out from a course or being admitted to a program, or to identify topics in written assignments. Popenici and Kerr (2017) define machine learning "as a subfield of artificial intelligence that includes software able to recognise patterns, make predictions, and apply newly discovered patterns to situations that were not included or covered by their initial design" (p. 2).

The concept of rational agents is central to AI: "An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators" (Russel & Norvig, 2010, p. 34). The vacuum-cleaner robot is a very simple form of an intelligent agent, but things become very complex and open-ended when we think about an automated taxi.

Experts in the field distinguish between weak and strong AI (see Russel & Norvig, 2010, p. 1020) or narrow and general AI (see Baker & Smith, 2019, p. 10). A philosophical question remains whether machines will be able to actually think or even develop consciousness in the future, rather than just simulating thinking and showing rational behaviour. It is unlikely that such strong or general AI will exist in the near future. We are therefore dealing here with GOFAI ("*good old-fashioned AI*", a term coined by the philosopher John Haugeland, 1985) in higher education – in the sense of agents and information systems that act *as if* they were intelligent.

Given this understanding of AI, what are potential areas of AI applications in education, and higher education in particular? Luckin, Holmes, Griffiths, and Forcier (2016) describe three categories of AI software applications in education that are available today: a) personal tutors, b) intelligent support for collaborative learning, and c) intelligent virtual reality.

Intelligent tutoring systems (ITS) can be used to simulate one-to-one personal tutoring. Based on learner models, algorithms and neural networks, they can make decisions about the learning path of an individual student and the content to select, provide cognitive scaffolding and help, to engage the student in dialogue. ITS have enormous potential, especially in large-scale distance teaching institutions, which run modules with thousands of students, where human one-to-one tutoring is impossible. A vast array of research shows that learning is a social exercise; interaction and collaboration are at the heart of the learning process (see for example Jonassen, Davidson, Collins, Campbell, & Haag, 1995). However, online collaboration has to be facilitated and moderated (Salmon, 2000). AIEd can contribute to collaborative learning by supporting adaptive group formation based on learner models, by facilitating online group interaction or by summarising discussions that can be used by a human tutor to guide students towards the aims and objectives of a course. Finally, also drawing on ITS, intelligent virtual reality (IVR) is used to engage and guide students in authentic virtual reality and game-based learning environments. Virtual agents can act as teachers, facilitators or students' peers, for example, in virtual or remote labs (Perez et al., 2017).

With the advancement of AIEd and the availability of (big) student data and learning analytics, Luckin et al. (2016) claim a "[r] enaissance in assessment" (p. 35). AI can provide just-in-time feedback and assessment. Rather than stop-and-test, AIEd can be built into learning activities for an ongoing analysis of student achievement. Algorithms have been used to predict the probability of a student failing an assignment or dropping out of a course with high levels of accuracy (e.g. Bahadır, 2016).

In their recent report, Baker and Smith (2019) approach educational AI tools from three different perspectives; a) learner-facing, b) teacher-facing, and c) system-facing AIEd. Learner-facing AI tools are software that students use to learn a subject matter, i.e. adaptive or personalised learning management systems or ITS. Teacher-facing systems are used to support the teacher and reduce his or her workload by automating tasks such as administration, assessment, feedback and plagiarism detection. AIEd tools also provide insight into the learning progress of students so that the teacher can proactively offer support and guidance where needed. System-facing AIEd are tools that provide information for administrators and managers on the institutional level, for example to monitor attrition patterns across faculties or colleges.

In the context of higher education, we use the concept of the student life-cycle (see Reid, 1995) as a framework to describe the various AI based services on the broader

institutional and administrative level, as well as for supporting the academic teaching and learning process in the narrower sense.

Method

The purpose of a systematic review is to answer specific questions, based on an explicit, systematic and replicable search strategy, with inclusion and exclusion criteria identifying studies to be included or excluded (Gough, Oliver & Thomas, 2017). Data is then coded and extracted from included studies, in order to synthesise findings and to shine light on their application in practice, as well as on gaps or contradictions. This contribution maps 146 articles on the topic of artificial intelligence in higher education.

Search strategy

The initial search string (see Table 1) and criteria (see Table 2) for this systematic review included peer-reviewed articles in English, reporting on artificial intelligence within education at any level, and indexed in three international databases; EBSCO Education Source, Web of Science and Scopus (covering titles, abstracts, and keywords). Whilst there are concerns about peer-review processes within the scientific community (e.g., Smith, 2006), articles in this review were limited to those published in peer-reviewed journals, due to their general trustworthiness in academia and the rigorous review processes undertaken (Nicholas et al., 2015). The search was undertaken in November 2018, with an initial 2656 records identified.

   After duplicates were removed, it was decided to limit articles to those published during or after 2007, as this was the year that iPhone's Siri was introduced; an algorithm-based personal assistant, started as an artificial intelligence project funded by the US Defense Advanced Research Projects Agency (DARPA) in 2001, turned into a company that was acquired by Apple Inc. It was also decided that the corpus would be limited to articles discussing applications of artificial intelligence in higher education only.

Screening and inter-rater reliability

The screening of 1549 titles and abstracts was carried out by a team of three coders and at this first screening stage, there was a requirement of sensitivity rather than specificity, i.e. papers were included rather than excluded. In order to reach consensus, the reasons for inclusion and exclusion for the first 80 articles were discussed at regular

Table 1 Initial search string

Topic	Search terms
Artificial intelligence	"artificial intelligence" OR "machine intelligence" OR "intelligent support" OR "intelligent virtual reality" OR "chat bot*" OR "machine learning" OR "automated tutor" OR "personal tutor*" OR "intelligent agent*" OR "expert system" OR "neural network" OR "natural language processing"
AND	
Education level	"higher education" OR college* OR undergrad* OR graduate OR postgrad* OR "K-12" OR kindergarten* OR "corporate training*" OR "professional training*" OR "primary school*" OR "middle school*" OR "high school*" OR "elementary school*" OR "vocational education" OR "adult education"
AND	
Learning setting	learn* OR student*

Table 2 Final inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Published 2007 – Nov 2018	Published before 2007
English or Spanish language	Not in English or Spanish
Higher education	Not higher education
Empirical, primary research	Not primary research (e.g., review)
Indexed in Web of Science, Scopus or EBSCO Education Source	Not a journal article
	No artificial intelligence
Artificial intelligence use in education	No learning setting

meetings. Twenty articles were randomly selected to evaluate the coding decisions of the three coders (A, B and C) to determine inter-rater reliability using Cohen's kappa (κ) (Cohen, 1960), which is a coefficient for the degree of consistency among raters, based on the number of codes in the coding scheme (Neumann, 2007, p. 326). Kappa values of .40−.60 are characterised as fair, .60 to .75 as good, and over .75 as excellent (Bakeman & Gottman, 1997; Fleiss, 1981). Coding consistency for inclusion or exclusion of articles between rater A and B was κ = .79, between rater A and C it was κ = .89, and between rater B and C it was κ = .69 (median = .79). Therefore, inter-rater reliability can be considered as excellent for the coding of inclusion and exclusion criteria.

After initial screening, 332 potential articles remained for screening on full text (see Fig. 1). However, 41 articles could not be retrieved, either through the library order scheme or by contacting authors. Therefore, 291 articles were retrieved, screened and coded, and following the exclusion of 149 papers, 146 articles remained for synthesis.[5]
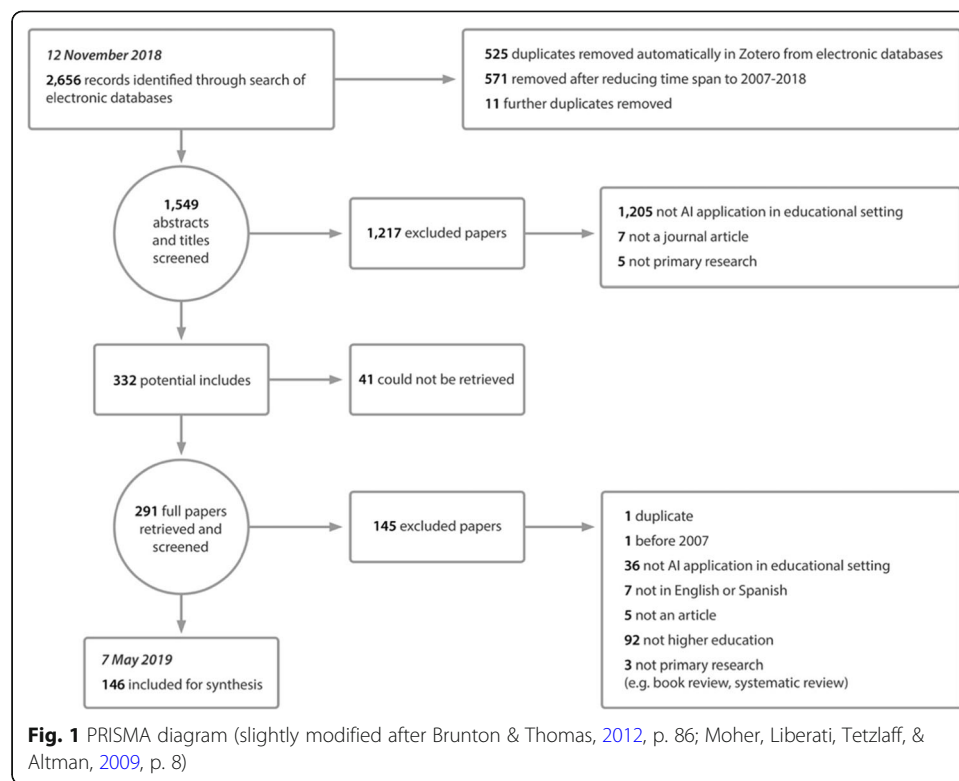
Coding, data extraction and analysis

In order to extract the data, all articles were uploaded into systematic review software EPPI Reviewer[6] and a coding system was developed. Codes included article information (year of publication, journal name, countries of authorship, discipline of first author), study design and execution (empirical or descriptive, educational setting) and how artificial intelligence was used (applications in the student life cycle, specific applications and methods). Articles were also coded on whether challenges and benefits of AI were present, and whether AI was defined. Descriptive data analysis was carried out with the statistics software R using the *tidyr* package (Wickham & Grolemund, 2016).

Limitations

Whilst this systematic review was undertaken as rigorously as possible, each review is limited by its search strategy. Although the three educational research databases chosen are large and international in scope, by applying the criteria of peer-reviewed articles published only in English or Spanish, research published on AI in other languages were not included in this review. This also applies to research in conference proceedings, book chapters or grey literature, or those articles not published in journals that are indexed in the three databases searched. In addition, although Spanish peer-reviewed articles were added according to inclusion criteria, no specific search string in the

---

[5]A file with all included references is available at: https://www.researchgate.net/publication/
335911716_AIED-Ref (CC-0; DOI: https://doi.org/10.13140/RG.2.2.13000.88321)
[6]https://eppi.ioe.ac.uk/cms/er4/ (accessed July 22, 2019)

Fig. 1 PRISMA diagram (slightly modified after Brunton & Thomas, 2012, p. 86; Moher, Liberati, Tetzlaff, & Altman, 2009, p. 8)

language was included, which narrows down the possibility of including Spanish papers that were not indexed with the chosen keywords. Future research could consider using a larger number of databases, publication types and publication languages, in order to widen the scope of the review. However, serious consideration would then need to be given to project resources and the manageability of the review (see Authors, in press).

Results

### Journals, authorship patterns and methods

Articles per year

There was a noticeable increase in the papers published from 2007 onwards. The number of included articles grew from six in 2007 to 23 in 2018 (see Fig. 2).

Journals

The papers included in the sample were published in 104 different journals. The greatest number of articles were published in the *International Journal of Artificial Intelligence in Education* (*n* = 11), followed by *Computers & Education* (*n* = 8), and the *International Journal of Emerging Technologies in Learning* (*n* = 5). Table 3 lists 19 journals that published at least two articles on AI in higher education from 2007 to 2018.

Countries

For the geographical distribution analysis of articles, the country of origin of the first author was taken into consideration (*n* = 38 countries). Table 4 shows 19 countries that

Fig. 2 Number of included articles per year (*n* = 146)
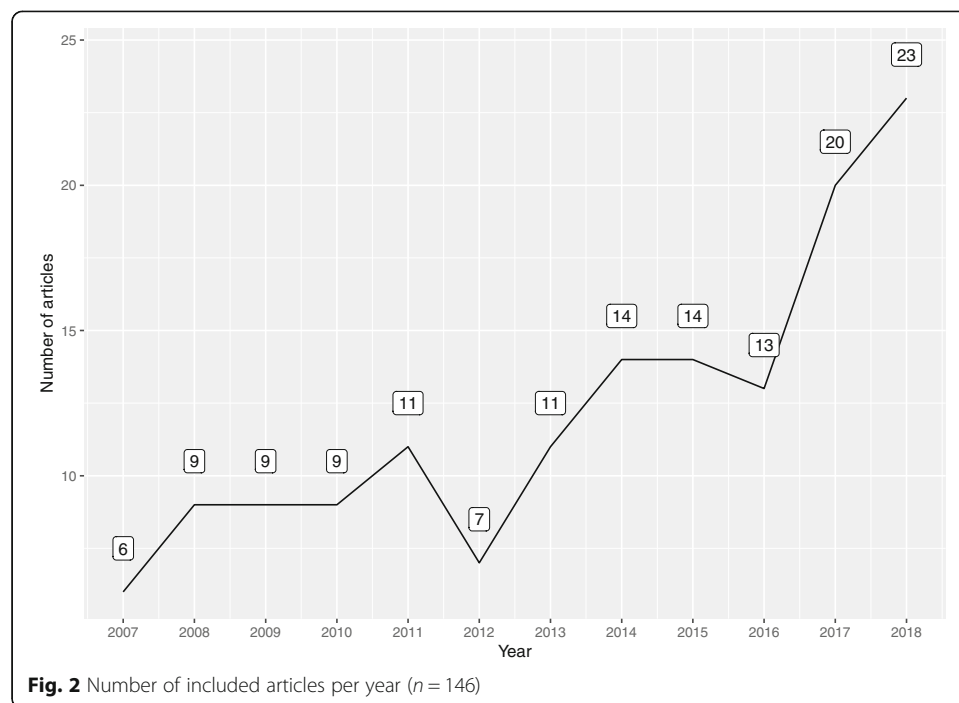
Table 3 Number of included articles (*n* = 146) by journal

Rank	Journal	*n*
1	International Journal of Artificial Intelligence in Education	11
2	Computers & Education	8
3	International Journal of Emerging Technologies in Learning	5
4	Decision Support Systems	3
	Expert Systems with Applications	3
	IEEE Transactions on Learning Technologies	3
	International Journal of Engineering and Technology	3
5	AI Magazine	2
	Computer Assisted Language Instruction Consortium Journal (CALICO)	2
	Computers in Human Behaviour	2
	Decision Sciences Journal of Innovative Education	2
	Educational Technology and Society	2
	Electronic Journal of e-Learning	2
	IEEE Journal on Selected Topics in Signal Processing	2
	International Journal of Engineering Education	2
	Journal of College Student Retention: Research, Theory and Practice	2
	Journal of Computer Assisted Learning	2
	Journal of Science Education & Technology	2
	Journal of the Learning Sciences	2
	Journal of Theoretical and Applied Information Technology	2
	others with one article	84

Table 4 Distribution of articles by country and cumulative percentages (*n* = 146)

Rank	Country	*n*	%	Rank	Country	*n*	%
1	USA	43	0.29		Malaysia	4	0.75
2	China	11	0.37	9	Iran	3	0.77
3	Taiwan	10	0.44		Mexico	3	0.79
4	Turkey	9	0.50	10	Argentina	2	0.80
5	UK	7	0.55		Croatia	2	0.82
6	India	6	0.59		Ecuador	2	0.83
	Spain	6	0.63		Germany	2	0.84
7	Canada	5	0.66		Indonesia	2	0.86
8	Australia	4	0.69		Singapore	2	0.87
	Greece	4	0.72		other	19	1.00

contributed at least two papers, and it reveals that 50% of all articles come from only four countries: USA, China, Taiwan, and Turkey.

Author affiliations

Again, the affiliation of the first author was taken into consideration (see Table 5). Researchers working in departments of Computer Science contributed by far the greatest number of papers (*n* = 61) followed by Science, Technology, Engineering and Mathematics (STEM) departments (*n* = 29). Only nine first authors came from an Education department, some reported dual affiliation with Education and Computer Science (*n* = 2), Education and Psychology (*n* = 1), or Education and STEM (*n* = 1).

Thus, 13 papers (8.9%) were written by first authors with an Education background. It is noticeable that three of them were contributed by researchers from the Teachers College at Columbia University, New York, USA (Baker, 2016; Paquette, Lebeau, Beaulieu, & Mayers, 2015; Perin & Lauterbach, 2018) – and they were all published in the same journal, i.e. the *International Journal of Artificial Intelligence in Education*.

Methods

Thirty studies (20.5%) were coded as being theoretical or descriptive in nature. The vast majority of studies (73.3%) applied quantitative methods, whilst only one (0.7%) was qualitative in nature and eight (5.5%) followed a mixed-methods approach. The purpose of the qualitative study, involving interviews with ESL students, was to explore the nature of written feedback coming from an automated essay scoring system compared to a human teacher (Dikli, 2010). In many cases, authors employed quasi-experimental methods, being an intentional sample divided into the experimental group, where an AI application (e.g. an intelligent tutoring system) was applied, and the control group without the intervention, followed by pre- and posttest (e.g. Adamson, Dyke, Jang, & Rosé, 2014).

Understanding of AI and critical reflection of challenges and risks

There are many different types and levels of AI mentioned in the articles, however only five out of 146 included articles (3.4%) provide an explicit definition of the term "Artificial Intelligence". The main characteristics of AI, described in all five studies, are the parallels between the human brain and artificial intelligence. The authors conceptualise

Table 5 Affiliation of the first author (*n* = 146 articles)

Affiliation	*n*	%
Computer Science	61	41.8
STEM	29	19.9
Arts, Humanities & Social Science	14	9.6
Education	9	6.2
Information Science	8	5.5
Psychology	8	5.5
Business & Law	6	4.1
not mentioned	4	2.7
Education and Computer Science	2	1.4
Computer Science and STEM	2	1.4
Computer Science and Arts Humanities & Social Science	1	0.7
Education and Psychology	1	0.7
Education and STEM	1	0.7
Total	146	100.0

AI as intelligent computer systems or intelligent agents with human features, such as the ability to memorise knowledge, to perceive and manipulate their environment in a similar way as humans, and to understand human natural language (see Huang, 2018; Lodhi, Mishra, Jain, & Bajaj, 2018; Welham, 2008). Dodigovic (2007) defines AI in her article as follows (p. 100):

> Artificial intelligence (AI) is a term referring to machines which emulate the behaviour of intelligent beings […] AI is an interdisciplinary area of knowledge and research, whose aim is to understand how the human mind works and how to apply the same principles in technology design. In language learning and teaching tasks, AI can be used to emulate the behaviour of a teacher or a learner […] . (p. 100)

Dodigovic is the only author who gives a definition of AI, and comes from an Arts, Humanities and Social Science department, taking into account aspects of AI and intelligent tutors in second language learning.

A stunningly low number of authors, only two out of 146 articles (1.4%), critically reflect upon ethical implications, challenges and risks of applying AI in education. Li (2007) deals with privacy concerns in his article about intelligent agent supported online learning:

> Privacy is also an important concern in applying agent-based personalised education. As discussed above, agents can autonomously learn many of students' personal information, like learning style and learning capability. In fact, personal information is private. Many students do not want others to know their private information, such as learning styles and/or capabilities. Students might show concern over possible discrimination from instructors in reference to learning performance due to special learning needs. Therefore, the privacy issue must be resolved before applying agent-based personalised teaching and learning technologies. (p. 327)

Another challenge of applying AI is mentioned by Welham (2008, p. 295) concerning the costs and time involved in developing and introducing AI-based methods that many public educational institutions cannot afford.

AI applications in higher education

As mentioned before, we used the concept of the student life-cycle (see Reid, 1995) as a framework to describe the various AI based services at the institutional and administrative level (e.g. admission, counselling, library services), as well as at the academic support level for teaching and learning (e.g. assessment, feedback, tutoring). Ninety-two studies (63.0%) were coded as relating to academic support services and 48 (32.8%) as administrative and institutional services; six studies (4.1%) covered both levels. The majority of studies addressed undergraduate students ($n = 91$, 62.3%) compared to 11 (7.5%) focussing on postgraduate students, and another 44 (30.1%) that did not specify the study level.

The iterative coding process led to the following four areas of AI applications with 17 sub-categories, covered in the publications: a) adaptive systems and personalisation, b) assessment and evaluation, c) profiling and prediction, and d) intelligent tutoring systems. Some studies addressed AI applications in more than one area (see Table 6).

The nature and scope of the various AI applications in higher education will be described along the lines of these four application categories in the following synthesis.

Profiling and prediction

The basis for many AI applications are learner models or profiles that allow prediction, for example of the likelihood of a student dropping out of a course or being admitted to a programme, in order to offer timely support or to provide feedback and guidance in content related matters throughout the learning process. Classification, modelling and prediction are an essential part of educational data mining (Phani Krishna, Mani Kumar, & Aruna Sri, 2018).

Most of the articles (55.2%, $n = 32$) address issues related to the institutional and administrative level, many (36.2%, $n = 21$) are related to academic teaching and learning at the course level, and five (8.6%) are concerned with both levels. Articles dealing with profiling and prediction were classified into three sub-categories; admission decisions and course scheduling ($n = 7$), drop-out and retention ($n = 23$), and student models and academic achievement ($n = 27$). One study that does not fall into any of these categories is the study by Ge and Xie (2015), which is concerned with forecasting the costs of a Chinese university to support management decisions based on an artificial neural network.

Table 6 Number of AI applications across studies ($n = 146$), multiple mentions possible

AI applications	*n*
Profiling and prediction (admissions decisions and course scheduling; drop-out and retention; student models and academic achievement)	58
Intelligent tutoring systems (teaching course content; diagnosing strengths and automated feedback; curating learning materials; facilitating collaboration; the teacher's perspective)	29
Assessment and evaluation (automated grading; feedback; evaluation of student understanding, engagement and academic integrity; evaluation of teaching)	36
Adaptive systems and personalisation (teaching course content; recommending personalized content; supporting teachers and learning design; using academic data to monitor and guide students; representation of knowledge in concept maps)	27
Total	150

All of the 58 studies in this area applied machine learning methods, to recognise and classify patterns, and to model student profiles to make predictions. Thus, they are all quantitative in nature. Many studies applied several machine learning algorithms (e.g. ANN, SVM, RF, NB; see Table 7)[7] and compared their overall prediction accuracy with conventional logistic regression. Table 7 shows that machine learning methods outper-formed logistic regression in all studies in terms of their classification accuracy in percent. To evaluate the performance of classifiers, the F1-score can also be used, which takes into account the number of positive instances correctly classified as positive, the number of negative instances incorrectly classified as positive, and the number of positive instances incorrectly classified as negative (Umer et al., 2017; for a brief overview of mea-sures of diagnostic accuracy, see Šimundić, 2009). The F1-score ranges between 0 and 1 with its best value at 1 (perfect precision and recall). Yoo and Kim (2014) reported high F1-scores of 0.848, 0.911, and 0.914 for J48, NB, and SVM, in a study to predict student's group project performance from online discussion participation.

Admission decisions and course scheduling Chen and Do (2014) point out that "the accurate prediction of students' academic performance is of importance for making ad-mission decisions as well as providing better educational services" (p. 18). Four studies aimed to predict whether or not a prospective student would be admitted to university. For example, Acikkar and Akay (2009) selected candidates for a School of Physical Edu-cation and Sports in Turkey based on a physical ability test, their scores in the National Selection and Placement Examination, and their graduation grade point average (GPA). They used the support vector machine (SVM) technique to classify the students and where able to predict admission decisions on a level of accuracy of 97.17% in 2006 and 90.51% in 2007. SVM was also applied by Andris, Cowen, and Wittenbach (2013) to find spatial patterns that might favour prospective college students from certain geo-graphic regions in the USA. Feng, Zhou, and Liu (2011) analysed enrolment data from 25 Chinese provinces as the training data to predict registration rates in other prov-inces using an artificial neural network (ANN) model. Machine learning methods and ANN are also used to predict student course selection behaviour to support course planning. Kardan, Sadeghi, Ghidary, and Sani (2013) investigated factors influencing student course selection, such as course and instructor characteristics, workload, mode of delivery and examination time, to develop a model to predict course selection with an ANN in two Computer Engineering and Information Technology Masters programs. In another paper from the same author team, a decision support system for course offer-ings was proposed (Kardan & Sadeghi, 2013). Overall, the research shows that admission decisions can be predicted at high levels of accuracy, so that an AI solution could relieves the administrative staff and allows them to focus on the more difficult cases.

Drop-out and retention Studies pertaining to drop-out and retention are intended to de-velop early warning systems to detect at-risk students in their first year (e.g., Alkhasawneh & Hargraves, 2014; Aluko, Adenuga, Kukoyi, Soyingbe, & Oyedeji, 2016; Hoffait & Schyns, 2017; Howard, Meehan, & Parnell, 2018) or to predict the attrition of undergraduate

---

[7]It is beyond the scope of this article to discuss the various machine learning methods for classification and prediction. Readers are therefore encouraged to refer to the literature referenced in the articles that are included in this review (e.g. Delen, 2010 and Umer, Susnjak, Mathrani, & Suriadi, 2017).

Table 7 Average performance of machine learning algorithms and logistic regression (%)

Author(s)	Prediction	ANN	RF/DT	SVM	NB	J48	LR
Acikkar and Akay (2009)	a)	–	–	93.8	–	–	–
Bahadır (2016)	b)	93.0	–	–	–	–	90.8
Delen (2010)	c)	86.5	87.2	87.2	–	–	86.1
Delen (2011)	c)	81.2	78.3	–	–	–	74.3
Hussain, Zhu, Zhang, and Abidi (2018)	d)	–	85.9	–	82.9	88.5	–
Oztekin (2016)	c)	71.6	73.8	77.6	–	–	–
Sreenivasa Rao, Swapna, and Praveen Kumar (2018)	a)	–	100.0	–	61.1	88.9	–
Teshnizi et al. (2015)	c)	84.3	–	–	–	–	77.5

Algorithms: *ANN* Artificial neural network, *RF/DT* Random forest / decision tree, *SVM* Support vector machines, *NB* Naïve Bayes, *J48* C4.5 decision tree, *LR* Logistic regression; Predictions: a) admission decisions; b) student academic performance; c) drop-out undergraduate students; d) student engagement

students in general (e.g., Oztekin, 2016; Raju & Schumacker, 2015). Delen (2011) used institutional data from 25,224 students enrolled as Freshmen in an American university over 8 years. In this study, three classification techniques were used to predict dropout: ANN, decision trees (DT) and logistic regression. The data contained variables related to students' demographic, academic, and financial characteristics (e.g. age, sex, ethnicity, GPA, TOEFL score, financial aid, student loan, etc.). Based on a 10-fold cross validation, Delen (2011) found that the ANN model worked best with an accuracy rate of 81.19% (see Table 7) and he concluded that the most important predictors of student drop-out are related to the student's past and present academic achievement, and whether they receive financial support. Sultana, Khan, and Abbas (2017, p. 107) discussed the impact of cognitive and non-cognitive features of students for predicting academic performance of undergraduate engineering students. In contrast to many other studies, they focused on non-cognitive variables to improve prediction accuracy, i.e. time management, self-concept, self-appraisal, leadership, and community support.

Student models and academic achievement Many more studies are concerned with profiling students and modelling learning behaviour to predict their academic achievements at the course level. Hussain et al. (2018) applied several machine learning algorithms to analyse student behavioural data from the virtual learning environment at the Open University UK, in order to predict student engagement, which is of particular importance at a large scale distance teaching university, where it is not possible to engage the majority of students in face-to-face sessions. The authors aim to develop an intelligent predictive system that enables instructors to automatically identify low-engaged students and then to make an intervention. Spikol, Ruffaldi, Dabisias, and Cukurova (2018) used face and hand tracking in workshops with engineering students to estimate success in project-based learning. They concluded that results generated from multimodal data can be used to inform teachers about key features of project-based learning activities. Blikstein et al. (2014) investigated patterns of how undergraduate students learn computer programming, based on over 150,000 code transcripts that the students created in software development projects. They found that their model, based on the process of programming, had better predictive power than the midterm grades.

Another example is the study of Babić ([2017](#)), who developed a model to predict student academic motivation based on their behaviour in an online learning environment.

The research on student models is an important foundation for the design of intelligent tutoring systems and adaptive learning environments.

Intelligent tutoring systems

All of the studies investigating intelligent tutoring systems (ITS) (*n* = 29) are only concerned with the teaching and learning level, except for one that is contextualised at the institutional and administrative level. The latter presents *StuA*, an interactive and intelligent student assistant that helps newcomers in a college by answering queries related to faculty members, examinations, extra curriculum activities, library services, etc. (Lodhi et al., [2018](#)).

The most common terms for referring to ITS described in the studies are *intelligent (online) tutors* or *intelligent tutoring systems* (e.g., in Dodigovic, [2007](#); Miwa, Terai, Kanzaki, & Nakaike, [2014](#)), although they are also identified often as *intelligent (software) agents* (e.g., Schiaffino, Garcia, & Amandi, [2008](#)), or *intelligent assistants* (e.g., in Casamayor, Amandi, & Campo, [2009](#); Jeschike, Jeschke, Pfeiffer, Reinhard, & Richter, [2007](#)). According to Welham ([2008](#)), the first ITS reported was the SCHOLAR system, launched in 1970, which allowed the reciprocal exchange of questions between teacher and student, but not holding a continuous conversation.

Huang and Chen ([2016](#), p. 341) describe the different models that are usually integrated in ITS: the student model (e.g. information about the student's knowledge level, cognitive ability, learning motivation, learning styles), the teacher model (e.g. analysis of the current state of students, select teaching strategies and methods, provide help and guidance), the domain model (knowledge representation of both students and teachers) and the diagnosis model (evaluation of errors and defects based on domain model).

The implementation and validation of the ITS presented in the studies usually took place over short-term periods (a course or a semester) and no longitudinal studies were identified, except for the study by Jackson and Cossitt ([2015](#)). On the other hand, most of the studies showed (sometimes slightly) positive / satisfactory preliminary results regarding the performance of the ITS, but they did not take into account the novelty effect that a new technological development could have in an educational context. One study presented negative results regarding the type of support that the ITS provided (Adamson et al., [2014](#)), which could have been more useful if it was more adjusted to the type of (in this case, more advanced) learners.

Overall, more research is needed on the effectiveness of ITS. The last meta-analysis of 39 ITS studies was published over 5 years ago: Steenbergen-Hu and Cooper ([2014](#)) found that ITS had a moderate effect of students' learning, and that ITS were less effective that human tutoring, but ITS outperformed all other instruction methods (such as traditional classroom instruction, reading printed or digital text, or homework assignments).

The studies addressing various ITS functions were classified as follows: teaching course content (*n* = 12), diagnosing strengths or gaps in students' knowledge and providing automated feedback (*n* = 7), curating learning materials based on students' needs (*n* = 3), and facilitating collaboration between learners (*n* = 2).

Teaching course content Most of the studies ($n = 4$) within this group focused on teaching Computer Science content (Dobre, 2014; Hooshyar, Ahmad, Yousefi, Yusop, & Horng, 2015; Howard, Jordan, di Eugenio, & Katz, 2017; Shen & Yang, 2011). Other studies included ITS teaching content for Mathematics (Miwa et al., 2014), Business Statistics and Accounting (Jackson & Cossitt, 2015; Palocsay & Stevens, 2008), Medicine (Payne et al., 2009) and writing and reading comprehension strategies for undergraduate Psychology students (Ray & Belden, 2007; Weston-Sementelli, Allen, & McNamara, 2018). Overall, these ITS focused on providing teaching content to students and, at the same time, supporting them by giving adaptive feedback and hints to solve questions related to the content, as well as detecting students' difficulties/errors when working with the content or the exercises. This is made possible by monitoring students' actions with the ITS.

In the study by Crown, Fuentes, Jones, Nambiar, and Crown (2011), a combination of teaching content through dialogue with a chatbot, that at the same time learns from this conversation - defined as a text-based conversational agent -, is described, which moves towards a more active, reflective and thinking student-centred learning approach. Duffy and Azevedo (2015) present an ITS called MetaTutor, which is designed to teach students about the human circulatory system, but it also puts emphasis on supporting students' self-regulatory processes assisted by the features included in the MetaTutor system (a timer, a toolbar to interact with different learning strategies, and learning goals, amongst others).

Diagnosing strengths or gaps in student knowledge, and providing automated feedback In most of the studies ($n = 4$) of this group, ITS are presented as a rather one-way communication from computer to student, concerning the gaps in students' knowledge and the provision of feedback. Three examples in the field of STEM have been found: two of them where the virtual assistance is presented as a feature in virtual laboratories by tutoring feedback and supervising student behaviour (Duarte, Butz, Miller, & Mahalingam, 2008; Ramírez, Rico, Riofrío-Luzcando, Berrocal-Lobo, & Antonio, 2018), and the third one is a stand-alone ITS in the field of Computer Science (Paquette et al., 2015). One study presents an ITS of this kind in the field of second language learning (Dodigovic, 2007).

In two studies, the function of diagnosing mistakes and the provision of feedback is accomplished by a dialogue between the student and the computer. For example, with an interactive ubiquitous teaching robot that bases its speech on question recognition (Umarani, Raviram, & Wahidabanu, 2011), or with the tutoring system, based on a tutorial dialogue toolkit for introductory college Physics (Chi, VanLehn, Litman, & Jordan, 2011). The same tutorial dialogue toolkit (TuTalk) is the core of the peer dialogue agent presented by Howard et al. (2017), where the ITS engages in a one-on-one problem-solving peer interaction with a student and can interact verbally, graphically and in a process-oriented way, and engage in collaborative problem solving instead of tutoring. This last study could be considered as part of a new category regarding peer-agent collaboration.

Curating learning materials based on student needs Two studies focused on this kind of ITS function (Jeschike et al., 2007; Schiaffino et al., 2008), and a third one

mentions it in a more descriptive way as a feature of the detection system presented (Hall Jr & Ko, 2008). Schiaffino et al. (2008) present eTeacher as a system for personalised assistance to e-learning students by observing their behaviour in the course and generating a student's profile. This enables the system to provide specific recommendations regarding the type of reading material and exercises done, as well as personalised courses of action. Jeschike et al. (2007) refers to an intelligent assistant contextualised in a virtual laboratory of statistical mechanics, where it presents exercises and the evaluation of the learners' input to content, and interactive course material that adapts to the learner.

Facilitating collaboration between learners Within this group we can identify only two studies: one focusing on supporting online collaborative learning discussions by using academically productive talk moves (Adamson et al., 2014); and the second one, on facilitating collaborative writing by providing automated feedback, generated automatic questions, and the analysis of the process (Calvo, O'Rourke, Jones, Yacef, & Reimann, 2011). Given the opportunities that the applications described in these studies afford for supporting collaboration among students, more research in this area would be desireable.

The teachers' perspective As mentioned above, Baker and Smith (2019, p.12) distinguish between student and teacher-facing AI. However, only two included articles in ITS focus on the teacher's perspective. Casamayor et al. (2009) focus on assisting teachers with the supervision and detection of conflictive cases in collaborative learning. In this study, the intelligent assistant provides the teachers with a summary of the individual progress of each group member and the type of participation each of them have had in their work groups, notification alerts derived from the detection of conflict situations, and information about the learning style of each student-logging interactions, so that the teachers can intervene when they consider it convenient. The other study put the emphasis on the ITS sharing teachers' tutoring tasks by providing immediate feedback (automating tasks), and leaving the teachers the role of providing new hints and the correct solution to the tasks (Chou, Huang, & Lin, 2011). The study of Chi et al. (2011) also mentions the ITS purpose to share teacher's tutoring tasks. The main aim in any of these cases is to reduce teacher's workload. Furthermore, many of the learner-facing studies deal with the teacher-facing functions too, although they do not put emphasis on the teacher's perspective.

Assessment and evaluation

Assessment and evaluation studies also largely focused on the level of teaching and learning (86%, $n = 31$), although five studies described applications at the institutional level. In order to gain an overview of student opinion about online and distance learning at their institution, academics at Anadolu University (Ozturk, Cicek, & Ergul, 2017) used sentiment analysis to analyse mentions by students on Twitter, using Twitter API Twython and terms relating to the system. This analysis of publicly accessible data, allowed researchers insight into student opinion, which otherwise may not have been accessible through their institutional LMS, and which can inform improvements to the

system. Two studies used AI to evaluate student Prior Learning and Recognition (PLAR); Kalz et al. (2008) used Latent Semantic Analysis and ePortfolios to inform personalised learning pathways for students, and Biletska, Biletskiy, Li, and Vovk (2010) used semantic web technologies to convert student credentials from different institutions, which could also provide information from course descriptions and topics, to allow for easier granting of credit. The final article at the institutional level (Sanchez et al., 2016) used an algorithm to match students to professional competencies and capabilities required by companies, in order to ensure alignment between courses and industry needs.

Overall, the studies show that AI applications can perform assessment and evaluation tasks at very high accuracy and efficiency levels. However, due to the need to calibrate and train the systems (supervised machine learning), they are more applicable to courses or programs with large student numbers.

Articles focusing on assessment and evaluation applications of AI at the teaching and learning level, were classified into four sub-categories; automated grading ($n = 13$), feedback ($n = 8$), evaluation of student understanding, engagement and academic integrity ($n = 5$), and evaluation of teaching ($n = 5$).

Automated grading Articles that utilised automated grading, or *Automated Essay Scoring* (AES) systems, came from a range of disciplines (e.g. Biology, Medicine, Business Studies, English as a Second Language), but were mostly focused on its use in undergraduate courses ($n = 10$), including those with low reading and writing ability (Perin & Lauterbach, 2018). Gierl, Latifi, Lai, Boulais, and Champlain's (2014) use of open source Java software LightSIDE to grade postgraduate medical student essays resulted in an agreement between the computer classification and human raters between 94.6% and 98.2%, which could enable reducing cost and the time associated with employing multiple human assessors for large-scale assessments (Barker, 2011; McNamara, Crossley, Roscoe, Allen, & Dai, 2015). However, they stressed that not all writing genres may be appropriate for AES and that it would be impractical to use in most small classrooms, due to the need to calibrate the system with a large number of pre-scored assessments. The benefits of using algorithms that find patterns in text responses, however, has been found to lead to encouraging more revisions by students (Ma & Slater, 2015) and to move away from merely measuring student knowledge and abilities by multiple choice tests (Nehm, Ha, & Mayfield, 2012). Continuing issues persist, however, in the quality of feedback provided by AES (Dikli, 2010), with Barker (2011) finding that the more detailed the feedback provided was, the more likely students were to question their grades, and a question was raised over the benefits of this feedback for beginning language students (Aluthman, 2016).

Feedback Articles concerned with feedback included a range of student-facing tools, including intelligent agents that provide students with prompts or guidance when they are confused or stalled in their work (Huang, Chen, Luo, Chen, & Chuang, 2008), software to alert trainee pilots when they are losing situation awareness whilst flying (Thatcher, 2014), and machine learning techniques with lexical features to generate automatic feedback and assist in improving student writing (Chodorow, Gamon, & Tetreault, 2010; Garcia-Gorrostieta, Lopez-Lopez, & Gonzalez-Lopez, 2018; Quixal & Meurers, 2016),

which can help reduce students cognitive overload (Yang, Wong, & Yeh, 2009). The automated feedback system based on adaptive testing reported by Barker (2010), for example, not only determines the most appropriate individual answers according to Bloom's cognitive levels, but also recommends additional materials and challenges.

Evaluation of student understanding, engagement and academic integrity Three articles reported on student-facing tools that evaluate student understanding of concepts (Jain, Gurupur, Schroeder, & Faulkenberry, 2014; Zhu, Marquez, & Yoo, 2015) and provide personalised assistance (Samarakou, Fylladitakis, Früh, Hatziapostolou, & Gelegenis, 2015). Hussain et al. (2018) used machine learning algorithms to evaluate student engagement in a social science course at the Open University, including final results, assessment scores and the number of clicks that students make in the VLE, which can alert instructors to the need for intervention, and Amigud, Arnedo-Moreno, Daradoumis, and Guerrero-Roldan (2017) used machine learning algorithms to check academic integrity, by assessing the likelihood of student work being similar to their other work. With a mean accuracy of 93%, this opens up possibilities of reducing the need for invigilators or to access student accounts, thereby reducing concerns surrounding privacy.

Evaluation of teaching Four studies used data mining algorithms to evaluate lecturer performance through course evaluations (Agaoglu, 2016; Ahmad & Rashid, 2016; DeCarlo & Rizk, 2010; Gutierrez, Canul-Reich, Ochoa Zezzatti, Margain, & Ponce, 2018), with Agaoglu (2016) finding, through using four different classification techniques, that many questions in the evaluation questionnaire were irrelevant. The application of an algorithm to evaluate the impact of teaching methods in a differential equations class, found that online homework with immediate feedback was more effective than clickers (Duzhin & Gustafsson, 2018). The study also found that, whilst previous exam results are generally good predictors for future exam results, they say very little about students' expected performance in project-based tasks.

Adaptive systems and personalisation Most of the studies on adaptive systems (85%, $n = 23$) are situated at the teaching and learning level, with four cases considering the institutional and administrative level. Two studies explored undergraduate students' academic advising (Alfarsi, Omar, & Alsinani, 2017; Feghali, Zbib, & Hallal, 2011), and Nguyen et al. (2018) focused on AI to support university career services. Ng, Wong, Lee, and Lee (2011) reported on the development of an agent-based distance LMS, designed to manage resources, support decision making and institutional policy, and assist with managing undergraduate student study flow (e.g. intake, exam and course management), by giving users access to data across disciplines, rather than just individual faculty areas.

There does not seem to be agreement within the studies on a common term for adaptive systems, and that is probably due to the diverse functions they carry out, which also supports the classification of studies. Some of those terms coincide in part with the ones used for ITS, e.g. *intelligent agents* (Li, 2007; Ng et al., 2011). The most general terms used are *intelligent e-learning system* (Kose & Arslan, 2016), *adaptive*

web-based learning system (Lo, Chan, & Yeh, 2012), or *intelligent teaching system* (Yuanyuan & Yajuan, 2014). As in ITS, most of the studies either describe the system or include a pilot study but no longer-term results are reported. Results from these pilot studies are usually reported as positive, except in Vlugter, Knott, McDonald, and Hall (2009), where the experimental group that used the dialogue-based computer assisted language-system scored lower than the control group in the delayed post-tests.

The 23 studies focused on teaching and learning can be classified into five sub-categories; teaching course content ($n = 7$), recommending/providing personalised content ($n = 5$), supporting teachers in learning and teaching design ($n = 3$), using academic data to monitor and guide students ($n = 2$), and supporting representation of knowledge using concept maps ($n = 2$). However, some studies were difficult to classify, due to their specific and unique functions; helping to organise online learning groups with similar interests (Yang, Wang, Shen, & Han, 2007), supporting business decisions through simulation (Ben-Zvi, 2012), or supporting changes in attitude and behaviour for patients with Anorexia Nervosa, through embodied conversational agents (Sebastian & Richards, 2017). Aparicio et al. (2018) present a study where no adaptive system application was analysed, rather students' perceptions of the use of information systems in education in general - and biomedical education in particular - were analysed, including *intelligent information access systems*.

Teaching course content The disciplines that are taught through adaptive systems are diverse, including environmental education (Huang, 2018), animation design (Yuanyuan & Yajuan, 2014), language learning (Jia, 2009; Vlugter et al., 2009), Computer Science (Iglesias, Martinez, Aler, & Fernandez, 2009) and Biology (Chaudhri et al., 2013). Walsh, Tamjidul, and Williams (2017), however, present an adaptive system based on machine learning-human machine learning symbiosis from a descriptive perspective, without specifying any discipline.

Recommending/providing personalised content This group refers to adaptive systems that deliver customised content, materials and exercises according to students' behaviour profiling in Business and Administration studies (Hall Jr & Ko, 2008) and Computer Science (Kose & Arslan, 2016; Lo et al., 2012). On the other hand, Tai, Wu, and Li (2008) present an e-learning recommendation system for online students to help them choose among courses, and Torres-Díaz, Infante Moro, and Valdiviezo Díaz (2014) emphasise the usefulness of (adaptive) recommendation systems in MOOCs to suggest actions, new items and users, according to students' personal preferences.

Supporting teachers in learning and teaching design In this group, three studies were identified. One study puts the emphasis on a hybrid recommender system of pedagogical patterns, to help teachers define their teaching strategies, according to the context of a specific class (Cobos et al., 2013), and another study presents a description of a metadata-based model to implement automatic learning designs that can solve detected problems (Camacho & Moreno, 2007). Li's (2007) descriptive study argues that intelligent agents save time for online instructors, by leaving the most repetitive tasks to the systems, so that they can focus more on creative work.

Using academic data to monitor and guide students The adaptive systems within this category focus on the extraction of student academic information to perform diagnostic tasks, and help tutors to offer a more proactive personal guidance (Rovira, Puertas, & Igual, 2017); or, in addition to that task, include performance evaluation and personalised assistance and feedback, such as the *Learner Diagnosis, Assistance, and Evaluation System based on AI (StuDiAsE)* for engineering learners (Samarakou et al., 2015).

Supporting representation of knowledge in concept maps To help build students' self-awareness of conceptual structures, concept maps can be quite useful. In the two studies of this group, an expert system was included, e.g. in order to accommodate selected peer ideas in the integrated concept maps and allow teachers to flexibly determine in which ways the selected concept maps are to be merged (*ICMSys*) (Kao, Chen, & Sun, 2010), or to help English as a Foreign Language college students to develop their reading comprehension through mental maps of referential identification (Yang et al., 2009). This latter system also includes system-guided instruction, practice and feedback.

Conclusions and implications for further educational research

In this paper, we have explored the field of AIEd research in terms of authorship and publication patterns. It is evident that US-American, Chinese, Taiwanese and Turkish colleagues (accounting for 50% of the publications as first authors) from Computer Science and STEM departments (62%) dominate the field. The leading journals are the *International Journal of Artificial Intelligence in Education, Computers & Education*, and the *International Journal of Emerging Technologies in Learning*.

More importantly, this study has provided an overview of the vast array of potential AI applications in higher education to support students, faculty members, and administrators. They were described in four broad areas (profiling and prediction, intelligent tutoring systems, assessment and evaluation, and adaptive systems and personalisation) with 17 sub-categories. This structure, which was derived from the systematic review, contributes to the understanding and conceptualisation of AIEd practice and research.

On the other hand, the lack of longitudinal studies and the substantial presence of descriptive and pilot studies from the technological perspective, as well as the prevalence of quantitative methods - especially quasi-experimental methods - in empirical studies, shows that there is still substantial room for educators to aim at innovative and meaningful research and practice with AIEd that could have learning impact within higher education, e.g. adopting design-based approaches (Easterday, Rees Lewis, & Gerber, 2018). A recent systematic literature review on personalisation in educational technology coincided with the predominance of experiences in technological developments, which also often used quantitative methods (Bartolomé, Castañeda, & Adell, 2018). Misiejuk and Wasson (2017, p. 61) noted in their systematic review on Learning Analytics that "there are very few implementation studies and impact studies" (p. 61), which is also similar to the findings in the present article.

The full consequences of AI development cannot yet be foreseen today, but it seems likely that AI applications will be a top educational technology issue for the next 20 years. AI-based tools and services have a high potential to support students, faculty members and administrators throughout the student lifecycle. The applications that are

described in this article provide enormous pedagogical opportunities for the design of intelligent student support systems, and for scaffolding student learning in adaptive and personalized learning environments. This applies in particular to large higher education institutions (such as open and distance teaching universities), where AIEd might help to overcome the dilemma of providing access to higher education for very large numbers of students (mass higher education). On the other hand, it might also help them to offer flexible, but also interactive and personalized learning opportunities, for example by relieving teachers from burdens, such as grading hundreds or even thousands of assignments, so that they can focus on their real task: empathic human teaching.

It is crucial to emphasise that educational technology is not (only) about technology – it is the pedagogical, ethical, social, cultural and economic dimensions of AIEd we should be concerned about. Selwyn (2016, p. 106) writes:

> The danger, of course, lies in seeing data and coding as an absolute rather than relative source of guidance and support. Education is far too complex to be reduced solely to data analysis and algorithms. As with digital technologies in general, digital data do not offer a neat technical fix to education dilemmas – no matter how compelling the output might be.

We should not strive for what is technically possible, but always ask ourselves what makes pedagogical sense. In China, systems are already being used to monitor student participation and expressions via face recognition in classrooms (so called Intelligent Classroom Behavior Management System, Smart Campus[8]) and display them to the teacher on a dashboard. This is an example of educational surveillance, and it is highly questionable whether such systems provide real added value for a good teacher who should be able to capture the dynamics in a learning group (online and in an on-campus setting) and respond empathically and in a pedagogically meaningful way. In this sense, it is crucial to adopt an ethics of care (Prinsloo, 2017) to start thinking on how we are exploring the potential of algorithmic decision-making systems that are embedded in AIEd applications. Furthermore, we should also always remember that AI systems "first and foremost, require control by humans. Even the smartest AI systems can make very stupid mistakes. […] AI Systems are only as smart as the date used to train them" (Kaplan & Haenlein, 2019, p. 25). Some critical voices in educational technology remind us that we should go beyond the tools, and talk again about learning and pedagogy, as well as acknowledging the human aspects of digital technology use in education (Castañeda & Selwyn, 2018). The new UNESCO report on challenges and opportunities of AIEd for sustainable development deals with various areas, all of which have an important pedagogical, social and ethical dimension, e.g. ensuring inclusion and equity in AIEd, preparing teachers for AI-powered education, developing quality and inclusive data systems, or ethics and transparency in data collection, use and dissemination (Pedró, Subosa, Rivas, & Valverde, 2019).

That being said, a stunning result of this review is the dramatic lack of critical reflection of the pedagogical and ethical implications as well as risks of implementing AI applications in higher education. Concerning ethical implications, privacy issues were also

---

[8]https://www.businessinsider.de/china-school-facial-recognition-technology-2018-5?r=US&IR=T (accessed July 5, 2019)

noted to be rarely addressed in empirical studies in a recent systematic review on Learning Analytics (Misiejuk & Wasson, 2017). More research is needed from educators and learning designers on how to integrate AI applications throughout the student lifecycle, to harness the enormous opportunities that they afford for creating intelligent learning and teaching systems. The low presence of authors affiliated with Education departments identified in our systematic review is evidence of the need for educational perspectives on these technological developments.

The lack of theory might be a syndrome within the field of educational technology in general. In a recent study, Hew, Lan, Tang, Jia, and Lo (2019) found that more than 40% of articles in three top educational technology journals were wholly a-theoretical. The systematic review by Bartolomé et al. (2018) also revealed this lack of explicit pedagogical perspectives in the studies analysed. The majority of research included in this systematic review is merely focused on analysing and finding patterns in data to develop models, and to make predictions that inform student and teacher facing applications, or to support administrative decisions using mathematical theories and machine learning methods that were developed decades ago (see Russel & Norvig, 2010). This kind of research is now possible through the growth of computing power and the vast availability of big digital student data. However, at this stage, there is very little evidence for the advancement of pedagogical and psychological learning theories related to AI driven educational technology. It is an important implication of this systematic review, that researchers are encouraged to be explicit about the theories that underpin empirical studies about the development and implementation of AIEd projects, in order to expand research to a broader level, helping us to understand the reasons and mechanisms behind this dynamic development that will have an enormous impact on higher education institutions in the various areas we have covered in this review.

Authors' information
Dr. Olaf Zawacki-Richter is a Professor of Educational Technology in the Faculty of Education and Social Sciences at the University of Oldenburg in Germany. He is the Director of the Center for Open Education Research (COER) and the Center for Lifelong Learning (C3L).
Dr. Victoria I. Marín is a Post-doctoral Researcher in the Faculty of Education and Social Sciences / Center for Open Education Research (COER) at the University of Oldenburg in Germany.
Melissa Bond is a PhD candidate and Research Associate in the Faculty of Education and Social Sciences / Center for Open Education Research (COER) at the University of Oldenburg in Germany.
Franziska Gouverneur is a Masters student and Research Assistant in the Faculty of Education and Social Sciences / Center for Open Education Research (COER) at the University of Oldenburg in Germany.

Competing interests

References

Acikkar, M., & Akay, M. F. (2009). Support vector machines for predicting the admission decision of a candidate to the School of Physical Education and Sports at Cukurova University. *Expert Systems with Applications*, *36*(3 PART 2), 7228–7233. https://doi.org/10.1016/j.eswa.2008.09.007.

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education*, *24*(1), 92–124. https://doi.org/10.1007/s40593-013-0012-6.

Agaoglu, M. (2016). Predicting instructor performance using data mining techniques in higher education. *IEEE Access*, *4*, 2379–2387. https://doi.org/10.1109/ACCESS.2016.2568756.

Ahmad, H., & Rashid, T. (2016). Lecturer performance analysis using multiple classifiers. *Journal of Computer Science*, *12*(5), 255–264. https://doi.org/10.3844/fjcssp.2016.255.264.

Alfarsi, G. M. S., Omar, K. A. M., & Alsinani, M. J. (2017). A rule-based system for advising undergraduate students. *Journal of Theoretical and Applied Information Technology*, *95*(11) Retrieved from http://www.jatit.org.

Alkhasawneh, R., & Hargraves, R. H. (2014). Developing a hybrid model to predict student first year retention in STEM disciplines using machine learning techniques. *Journal of STEM Education: Innovations & Research*, *15*(3), 35–42 https://core.ac.uk/download/pdf/51289621.pdf.

Aluko, R. O., Adenuga, O. A., Kukoyi, P. O., Soyingbe, A. A., & Oyedeji, J. O. (2016). Predicting the academic success of architecture students by pre-enrolment requirement: Using machine-learning techniques. *Construction Economics and Building*, *16*(4), 86–98. https://doi.org/10.5130/AJCEB.v16i4.5184.

Aluthman, E. S. (2016). The effect of using automated essay evaluation on ESL undergraduate students' writing skill. *International Journal of English Linguistics*, *6*(5), 54–67. https://doi.org/10.5539/ijel.v6n5p54.

Amigud, A., Arnedo-Moreno, J., Daradoumis, T., & Guerrero-Roldan, A.-E. (2017). Using learning analytics for preserving academic integrity. *International Review of Research in Open and Distance Learning*, *18*(5), 192–210. https://doi.org/10.19173/irrodl.v18i5.3103.

Andris, C., Cowen, D., & Wittenbach, J. (2013). Support vector machine for spatial variation. *Transactions in GIS*, *17*(1), 41–61. https://doi.org/10.1111/j.1467-9671.2012.01354.x.

Aparicio, F., Morales-Botello, M. L., Rubio, M., Hernando, A., Muñoz, R., López-Fernández, H., … de Buenaga, M. (2018). Perceptions of the use of intelligent information access systems in university level active learning activities among teachers of biomedical subjects. *International Journal of Medical Informatics*, *112*(December 2017), 21–33. https://doi.org/10.1016/j.ijmedinf.2017.12.016.

Babić, I. D. (2017). Machine learning methods in predicting the student academic motivation. *Croatian Operational Research Review*, *8*(2), 443–461. https://doi.org/10.17535/crorr.2017.0028.

Bahadır, E. (2016). Using neural network and logistic regression analysis to predict prospective mathematics teachers' academic success upon entering graduate education. *Kuram ve Uygulamada Egitim Bilimleri*, *16*(3), 943–964. https://doi.org/10.12738/estp.2016.3.0214.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction - an introduction to sequential analysis*. Cambridge: Cambridge University Press.

Baker, R. S. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600–614. https://doi.org/10.1007/s40593-016-0105-0.

Baker, T., & Smith, L. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. Retrieved from Nesta Foundation website: https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf

Barker, T. (2010). An automated feedback system based on adaptive testing: Extending the model. *International Journal of Emerging Technologies in Learning*, *5*(2), 11–14. https://doi.org/10.3991/ijet.v5i2.1235.

Barker, T. (2011). An automated individual feedback and marking system: An empirical study. *Electronic Journal of E-Learning*, *9*(1), 1–14 https://www.learntechlib.org/p/52053/.

Bartolomé, A., Castañeda, L., & Adell, J. (2018). Personalisation in educational technology: The absence of underlying pedagogies. *International Journal of Educational Technology in Higher Education*, *15*(14). https://doi.org/10.1186/s41239-018-0095-0.

Ben-Zvi, T. (2012). Measuring the perceived effectiveness of decision support systems and their impact on performance. *Decision Support Systems*, *54*(1), 248–256. https://doi.org/10.1016/j.dss.2012.05.033.

Biletska, O., Biletskiy, Y., Li, H., & Vovk, R. (2010). A semantic approach to expert system for e-assessment of credentials and competencies. *Expert Systems with Applications*, *37*(10), 7003–7014. https://doi.org/10.1016/j.eswa.2010.03.018.

Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *Journal of the Learning Sciences*, *23*(4), 561–599. https://doi.org/10.1080/10508406.2014.954750.

Brunton, J., & Thomas, J. (2012). Information management in systematic reviews. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews*, (pp. 83–106). London: SAGE.

Calvo, R. A., O'Rourke, S. T., Jones, J., Yacef, K., & Reimann, P. (2011). Collaborative writing support tools on the cloud. *IEEE Transactions on Learning Technologies*, *4*(1), 88–97 https://www.learntechlib.org/p/73461/.

Camacho, D., & Moreno, M. D. R. (2007). Towards an automatic monitoring for higher education learning design. *International Journal of Metadata, Semantics and Ontologies*, *2*(1), 1. https://doi.org/10.1504/ijmso.2007.015071.

Casamayor, A., Amandi, A., & Campo, M. (2009). Intelligent assistance for teachers in collaborative e-learning environments. *Computers & Education*, *53*(4), 1147–1154. https://doi.org/10.1016/j.compedu.2009.05.025.

Castañeda, L., & Selwyn, N. (2018). More than tools? Making sense of he ongoing digitizations of higher education. *International Journal of Educational Technology in Higher Education*, *15*(22). https://doi.org/10.1186/s41239-018-0109-y.

Chaudhri, V. K., Cheng, B., Overtholtzer, A., Roschelle, J., Spaulding, A., Clark, P., … Gunning, D. (2013). Inquire biology: A textbook that answers questions. *AI Magazine*, *34*(3), 55–55. https://doi.org/10.1609/aimag.v34i3.2486.

Chen, J.-F., & Do, Q. H. (2014). Training neural networks to predict student academic performance: A comparison of cuckoo search and gravitational search algorithms. *International Journal of Computational Intelligence and Applications*, *13*(1). https://doi.org/10.1142/S1469026814500059.

Chi, M., VanLehn, K., Litman, D., & Jordan, P. (2011). Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, *21*(1), 137–180. https://doi.org/10.1007/s11257-010-9093-1.

Chodorow, M., Gamon, M., & Tetreault, J. (2010). The utility of article and preposition error correction systems for English language learners: Feedback and assessment. *Language Testing*, *27*(3), 419–436. https://doi.org/10.1177/0265532210364391.

Chou, C.-Y., Huang, B.-H., & Lin, C.-J. (2011). Complementary machine intelligence and human intelligence in virtual teaching assistant for tutoring program tracing. *Computers & Education*, *57*(4), 2303–2312 https://www.learntechlib.org/p/167322/.

Cobos, C., Rodriguez, O., Rivera, J., Betancourt, J., Mendoza, M., León, E., & Herrera-Viedma, E. (2013). A hybrid system of pedagogical pattern recommendations based on singular value decomposition and variable data attributes. *Information Processing and Management*, *49*(3), 607–625. https://doi.org/10.1016/j.ipm.2012.12.002.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46. https://doi.org/10.1177/001316446002000104.

Contact North. (2018). Ten facts about artificial intelligence in teaching and learning. Retrieved from https://teachonline.ca/sites/default/files/tools-trends/downloads/ten_facts_about_artificial_intelligence.pdf

Crown, S., Fuentes, A., Jones, R., Nambiar, R., & Crown, D. (2011). Anne G. Neering: Interactive chatbot to engage and motivate engineering students. *Computers in Education Journal*, *21*(2), 24–34.

DeCarlo, P., & Rizk, N. (2010). The design and development of an expert system prototype for enhancing exam quality. *International Journal of Advanced Corporate Learning*, *3*(3), 10–13. https://doi.org/10.3991/ijac.v3i3.1356.

Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems*, *49*(4), 498–506. https://doi.org/10.1016/j.dss.2010.06.003.

Delen, D. (2011). Predicting student attrition with data mining methods. *Journal of College Student Retention: Research, Theory and Practice*, *13*(1), 17–35. https://doi.org/10.2190/CS.13.1.b.

Dikli, S. (2010). The nature of automated essay scoring feedback. *CALICO Journal*, *28*(1), 99–134. https://doi.org/10.11139/cj.28.1.99-134.

Dobre, I. (2014). Assessing the student's knowledge in informatics discipline using the METEOR metric. *Mediterranean Journal of Social Sciences*, *5*(19), 84–92. https://doi.org/10.5901/mjss.2014.v5n19p84.

Dodigovic, M. (2007). Artificial intelligence and second language learning: An efficient approach to error remediation. *Language Awareness*, *16*(2), 99–113. https://doi.org/10.2167/la416.0.

Duarte, M., Butz, B., Miller, S., & Mahalingam, A. (2008). An intelligent universal virtual laboratory (UVL). *IEEE Transactions on Education*, *51*(1), 2–9. https://doi.org/10.1109/SSST.2002.1027009.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348. https://doi.org/10.1016/j.chb.2015.05.041.

Duzhin, F., & Gustafsson, A. (2018). Machine learning-based app for self-evaluation of teacher-specific instructional style and tools. *Education Sciences*, *8*(1). https://doi.org/10.3390/educsci8010007.

Easterday, M. W., Rees Lewis, D. G., & Gerber, E. M. (2018). The logic of design research. *Learning: Research and Practice*, *4*(2), 131–160. https://doi.org/10.1080/23735082.2017.1286367.

EDUCAUSE. (2018). Horizon report: 2018 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium website: https://library.educause.edu/~/media/files/library/2018/8/2018horizonreport.pdf

EDUCAUSE. (2019). Horizon report: 2019 higher education edition. Retrieved from EDUCAUSE Learning Initiative and The New Media Consortium website: https://library.educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf

Feghali, T., Zbib, I., & Hallal, S. (2011). A web-based decision support tool for academic advising. *Educational Technology and Society*, *14*(1), 82–94 https://www.learntechlib.org/p/52325/.

Feng, S., Zhou, S., & Liu, Y. (2011). Research on data mining in university admissions decision-making. *International Journal of Advancements in Computing Technology*, *3*(6), 176–186. https://doi.org/10.4156/ijact.vol3.issue6.21.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.

Garcia-Gorrostieta, J. M., Lopez-Lopez, A., & Gonzalez-Lopez, S. (2018). Automatic argument assessment of final project reports of computer engineering students. Computer Applications in Engineering Education, 26(5), 1217–1226. https://doi.org/10.1002/cae.21996

Ge, C., & Xie, J. (2015). Application of grey forecasting model based on improved residual correction in the cost estimation of university education. *International Journal of Emerging Technologies in Learning*, *10*(8), 30–33. https://doi.org/10.3991/ijet.v10i8.5215.

Gierl, M., Latifi, S., Lai, H., Boulais, A., & Champlain, A. (2014). Automated essay scoring and the future of educational assessment in medical education. *Medical Education*, *48*(10), 950–962. https://doi.org/10.1111/medu.12517.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*, (2nd ed.,). Los Angeles: SAGE.

Gutierrez, G., Canul-Reich, J., Ochoa Zezzatti, A., Margain, L., & Ponce, J. (2018). Mining: Students comments about teacher performance assessment using machine learning algorithms. *International Journal of Combinatorial Optimization Problems and Informatics*, *9*(3), 26–40 https://ijcopi.org/index.php/ojs/article/view/99.

Hall Jr., O. P., & Ko, K. (2008). Customized content delivery for graduate management education: Application to business statistics. *Journal of Statistics Education*, *16*(3). https://doi.org/10.1080/10691898.2008.11889571.

Haugeland, J. (1985). Artificial intelligence: The very idea. Cambridge, Mass.: MIT Press

Hew, K. F., Lan, M., Tang, Y., Jia, C., & Lo, C. K. (2019). Where is the "theory" within the field of educational technology research? *British Journal of Educational Technology*, *50*(3), 956–971. https://doi.org/10.1111/bjet.12770.

Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, *9*(1), 51. https://doi.org/10.3390/educsci9010051.

Hoffait, A.-S., & Schyns, M. (2017). Early detection of university students with potential difficulties. *Decision Support Systems*, *101*, 1–11. https://doi.org/10.1016/j.dss.2017.05.003.

Hooshyar, D., Ahmad, R., Yousefi, M., Yusop, F., & Horng, S. (2015). A flowchart-based intelligent tutoring system for improving problem-solving skills of novice programmers. *Journal of Computer Assisted Learning*, *31*(4), 345–361. https://doi.org/10.1111/jcal.12099.

Howard, C., Jordan, P., di Eugenio, B., & Katz, S. (2017). Shifting the load: A peer dialogue agent that encourages its human collaborator to contribute more to problem solving. *International Journal of Artificial Intelligence in Education*, *27*(1), 101–129. https://doi.org/10.1007/s40593-015-0071-y.

Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *Internet and Higher Education*, *37*, 66–75. https://doi.org/10.1016/j.iheduc.2018.02.001.

Huang, C.-J., Chen, C.-H., Luo, Y.-C., Chen, H.-X., & Chuang, Y.-T. (2008). Developing an intelligent diagnosis and assessment e-Learning tool for introductory programming. *Educational Technology & Society*, *11*(4), 139–157 https://www.jstor.org/stable/jeductechsoci.11.4.139.

Huang, J., & Chen, Z. (2016). The research and design of web-based intelligent tutoring system. *International Journal of Multimedia and Ubiquitous Engineering*, *11*(6), 337–348. https://doi.org/10.14257/ijmue.2016.11.6.30.

Huang, S. P. (2018). Effects of using artificial intelligence teaching system for environmental education on environmental knowledge and attitude. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(7), 3277–3284. https://doi.org/10.29333/ejmste/91248.

Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student engagement predictions in an e-Learning system and their impact on student course assessment scores. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2018/6347186.

Iglesias, A., Martinez, P., Aler, R., & Fernandez, F. (2009). Reinforcement learning of pedagogical policies in adaptive and intelligent educational systems. *Knowledge-Based Systems*, *22*(4), 266–270 https://e-archivo.uc3m.es/bitstream/handle/10016/6502/reinforcement_aler_KBS_2009_ps.pdf?sequence=1&isAllowed=y.

Jackson, M., & Cossitt, B. (2015). Is intelligent online tutoring software useful in refreshing financial accounting knowledge? *Advances in Accounting Education: Teaching and Curriculum Innovations*, *16*, 1–19. https://doi.org/10.1108/S1085-462220150000016001.

Jain, G. P., Gurupur, V. P., Schroeder, J. L., & Faulkenberry, E. D. (2014). Artificial intelligence-based student learning evaluation: A concept map-based approach for analyzing a student's understanding of a topic. *IEEE Transactions on Learning Technologies*, *7*(3), 267–279. https://doi.org/10.1109/TLT.2014.2330297.

Jeschike, M., Jeschke, S., Pfeiffer, O., Reinhard, R., & Richter, T. (2007). Equipping virtual laboratories with intelligent training scenarios. *AACE Journal*, *15*(4), 413–436 https://www.learntechlib.org/primary/p/23636/.

Jia, J. (2009). An AI framework to teach English as a foreign language: CSIEC. *AI Magazine*, *30*(2), 59–59. https://doi.org/10.1609/aimag.v30i2.2232.

Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. *American Journal of Distance Education*, *9*(2), 7–25. https://doi.org/10.1080/08923649509526885.

Kalz, M., van Bruggen, J., Giesbers, B., Waterink, W., Eshuis, J., & Koper, R. (2008). A model for new linkages for prior learning assessment. *Campus-Wide Information Systems*, *25*(4), 233–243. https://doi.org/10.1108/10650740810900676.

Kao, Chen, & Sun (2010). Using an e-Learning system with integrated concept maps to improve conceptual understanding. *International Journal of Instructional Media*, *37*(2), 151–151.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004.

Kardan, A. A., & Sadeghi, H. (2013). A decision support system for course offering in online higher education institutes. *International Journal of Computational Intelligence Systems*, *6*(5), 928–942. https://doi.org/10.1080/18756891.2013.808428.

Kardan, A. A., Sadeghi, H., Ghidary, S. S., & Sani, M. R. F. (2013). Prediction of student course selection in online higher education institutes using neural network. *Computers and Education*, *65*, 1–11. https://doi.org/10.1016/j.compedu.2013.01.015.

Kose, U., & Arslan, A. (2016). Intelligent e-Learning system for improving students' academic achievements in computer programming courses. *International Journal of Engineering Education*, *32*(1, A), 185–198.

Li, X. (2007). Intelligent agent-supported online education. *Decision Sciences Journal of Innovative Education*, *5*(2), 311–331. https://doi.org/10.1111/j.1540-4609.2007.00143.x.

Lo, J. J., Chan, Y. C., & Yeh, S. W. (2012). Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Computers and Education*, *58*(1), 209–222. https://doi.org/10.1016/j.compedu.2011.08.018.

Lodhi, P., Mishra, O., Jain, S., & Bajaj, V. (2018). StuA: An intelligent student assistant. *International Journal of Interactive Multimedia and Artificial Intelligence*, *5*(2), 17–25. https://doi.org/10.9781/ijimai.2018.02.008.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed - an argument for AI in education. Retrieved from http://discovery.ucl.ac.uk/1475756/

Ma, H., & Slater, T. (2015). Using the developmental path of cause to bridge the gap between AWE scores and writing teachers' evaluations. *Writing & Pedagogy*, *7*(2), 395–422. https://doi.org/10.1558/wap.v7i2-3.26376.

McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., & Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing*, *23*, 35–59. https://doi.org/10.1016/j.asw.2014.09.002.

Misiejuk, K., & Wasson, B. (2017). *State of the field report on learning analytics. SLATE report 2017–2*. Bergen: Centre for the Science of Learning & Technology (SLATE) Retrieved from http://bora.uib.no/handle/1956/17740.

Miwa, K., Terai, H., Kanzaki, N., & Nakaike, R. (2014). An intelligent tutoring system with variable levels of instructional support for instructing natural deduction. *Transactions of the Japanese Society for Artificial Intelligence*, *29*(1), 148–156. https://doi.org/10.1527/tjsai.29.148.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*, b2535. https://doi.org/10.1136/bmj.b2535 Clinical Research Ed.

Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, *21*(1), 183–196. https://doi.org/10.1007/s10956-011-9300-9.

Neumann, W. L. (2007). *Social research methods: Qualitative and quantitative approaches*. Boston: Pearson.

Ng, S. C., Wong, C. K., Lee, T. S., & Lee, F. Y. (2011). Design of an agent-based academic information system for effective education management. *Information Technology Journal*, *10*(9), 1784–1788. https://doi.org/10.3923/itj.2011.1784.1788.

Nguyen, J., Sánchez-Hernández, G., Armisen, A., Agell, N., Rovira, X., & Angulo, C. (2018). A linguistic multi-criteria decision-aiding system to support university career services. *Applied Soft Computing Journal*, *67*, 933–940. https://doi.org/10.1016/j.asoc.2017.06.052.

Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R., … Levine, K. (2015). Peer review: still king in the digital age. *Learned Publishing*, *28*(1), 15–21. https://doi.org/10.1087/20150104.

Oztekin, A. (2016). A hybrid data analytic approach to predict college graduation status and its determinative factors. *Industrial Management and Data Systems*, *116*(8), 1678–1699. https://doi.org/10.1108/IMDS-09-2015-0363.

Ozturk, Z. K., Cicek, Z. I. E., & Ergul, Z. (2017). Sentiment analysis: An application to Anadolu University. *Acta Physica Polonica A*, *132*(3), 753–755. https://doi.org/10.12693/APhysPolA.132.753.

Palocsay, S. W., & Stevens, S. P. (2008). A study of the effectiveness of web-based homework in teaching undergraduate business statistics. *Decision Sciences Journal of Innovative Education*, *6*(2), 213–232. https://doi.org/10.1111/j.1540-4609.2008.00167.x.

Paquette, L., Lebeau, J. F., Beaulieu, G., & Mayers, A. (2015). Designing a knowledge representation approach for the generation of pedagogical interventions by MTTs. *International Journal of Artificial Intelligence in Education*, *25*(1), 118–156 https://www.learntechlib.org/p/168275/.

Payne, V. L., Medvedeva, O., Legowski, E., Castine, M., Tseytlin, E., Jukic, D., & Crowley, R. S. (2009). Effect of a limited-enforcement intelligent tutoring system in dermatopathology on student errors, goals and solution paths. *Artificial Intelligence in Medicine*, *47*(3), 175–197. https://doi.org/10.1016/j.artmed.2009.07.002.

Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*. Paris: UNESCO.

Perez, S., Massey-Allard, J., Butler, D., Ives, J., Bonn, D., Yee, N., & Roll, I. (2017). Identifying productive inquiry in virtual labs using sequence mining. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial intelligence in education*, (vol. 10,331, pp. 287–298). https://doi.org/10.1007/978-3-319-61425-0_24.

Perin, D., & Lauterbach, M. (2018). Assessing text-based writing of low-skilled college students. *International Journal of Artificial Intelligence in Education*, *28*(1), 56–78. https://doi.org/10.1007/s40593-016-0122-z.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden; Oxford: Blackwell Pub.

Phani Krishna, K. V., Mani Kumar, M., & Aruna Sri, P. S. G. (2018). Student information system and performance retrieval through dashboard. *International Journal of Engineering and Technology (UAE)*, *7*, 682–685. https://doi.org/10.14419/ijet.v7i2.7.10922.

Popenici, S., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*. https://doi.org/10.1186/s41039-017-0062-8.

Prinsloo, P. (2017). Fleeing from Frankenstein's monster and meeting Kafka on the way: Algorithmic decision-making in higher education. *E-Learning and Digital Media*, *14*(3), 138–163. https://doi.org/10.1177/2042753017731355.

Quixal, M., & Meurers, D. (2016). How can writing tasks be characterized in a way serving pedagogical goals and automatic analysis needs? *Calico Journal*, *33*(1), 19–48. https://doi.org/10.1558/cj.v33i1.26543.

Raju, D., & Schumacker, R. (2015). Exploring student characteristics of retention that lead to graduation in higher education using data mining models. *Journal of College Student Retention: Research, Theory and Practice*, *16*(4), 563–591. https://doi.org/10.2190/CS.16.4.e.

Ramírez, J., Rico, M., Riofrío-Luzcando, D., Berrocal-Lobo, M., & Antonio, A. (2018). Students' evaluation of a virtual world for procedural training in a tertiary-education course. *Journal of Educational Computing Research*, *56*(1), 23–47. https://doi.org/10.1177/0735633117706047.

Ray, R. D., & Belden, N. (2007). Teaching college level content and reading comprehension skills simultaneously via an artificially intelligent adaptive computerized instructional system. *Psychological Record*, *57*(2), 201–218 https://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1103&context=tpr.

Reid, J. (1995). Managing learner support. In F. Lockwood (Ed.), *Open and distance learning today*, (pp. 265–275). London: Routledge.

Rovira, S., Puertas, E., & Igual, L. (2017). Data-driven system to predict academic grades and dropout. *PLoS One*, *12*(2), 1–21. https://doi.org/10.1371/journal.pone.0171207.

Russel, S., & Norvig, P. (2010). *Artificial intelligence - a modern approach*. New Jersey: Pearson Education.

Salmon, G. (2000). *E-moderating - the key to teaching and learning online*, (1st ed.,). London: Routledge.

Samarakou, M., Fylladitakis, E. D., Früh, W. G., Hatziapostolou, A., & Gelegenis, J. J. (2015). An advanced eLearning environment developed for engineering learners. *International Journal of Emerging Technologies in Learning*, *10*(3), 22–33. https://doi.org/10.3991/ijet.v10i3.4484.

Sanchez, E. L., Santos-Olmo, A., Alvarez, E., Huerta, M., Camacho, S., & Fernandez-Medina, E. (2016). Development of an expert system for the evaluation of students' curricula on the basis of competencies. *Future Internet*, *8*(2). https://doi.org/10.3390/fi8020022.

Schiaffino, S., Garcia, P., & Amandi, A. (2008). eTeacher: Providing personalized assistance to e-learning students. *Computers & Education*, *51*(4), 1744–1754. https://doi.org/10.1016/j.compedu.2008.05.008.

Sebastian, J., & Richards, D. (2017). Changing stigmatizing attitudes to mental health via education and contact with embodied conversational agents. *Computers in Human Behavior*, *73*, 479–488. https://doi.org/10.1016/j.chb.2017.03.071.

Selwyn, N. (2016). Is technology good for education? Cambridge, UK: Malden, MA : Polity Press.

Shen, V. R. L., & Yang, C.-Y. (2011). Intelligent multiagent tutoring system in artificial intelligence. *International Journal of Engineering Education*, *27*(2), 248–256.

Šimundić, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. *Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*, *19*(4), 203–2011 https://www.ncbi.nlm.nih.gov/pubmed/27683318.

Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, *99*, 178–182. https://doi.org/10.1258/jrsm.99.4.178.

Spikol, D., Ruffaldi, E., Dabisias, G., & Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*, *34*(4), 366–377. https://doi.org/10.1111/jcal.12263.

Sreenivasa Rao, K., Swapna, N., & Praveen Kumar, P. (2018). Educational data mining for student placement prediction using machine learning algorithms. *International Journal of Engineering and Technology (UAE)*, *7*(1.2), 43–46. https://doi.org/10.14419/ijet.v7i1.2.8988.

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, *106*(2), 331–347. https://doi.org/10.1037/a0034752.

Sultana, S., Khan, S., & Abbas, M. (2017). Predicting performance of electrical engineering students using cognitive and non-cognitive features for identification of potential dropouts. *International Journal of Electrical Engineering Education*, *54*(2), 105–118. https://doi.org/10.1177/0020720916688484.

Tai, D. W. S., Wu, H. J., & Li, P. H. (2008). Effective e-learning recommendation system based on self-organizing maps and association mining. *Electronic Library*, *26*(3), 329–344. https://doi.org/10.1108/02640470810879482.

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. London: Penguin Books.

Teshnizi, S. H., & Ayatollahi, S. M. T. (2015). A comparison of logistic regression model and artificial neural networks in predicting of student's academic failure. Acta Informatica Medica, 23(5), 296–300. https://doi.org/10.5455/aim.2015.23.296–300

Thatcher, S. J. (2014). The use of artificial intelligence in the learning of flight crew situation awareness in an undergraduate aviation programme. *World Transactions on Engineering and Technology Education*, *12*(4), 764–768 https://www.semanticscholar.org/paper/The-use-of-artificial-intelligence-in-the-learning-Thatcher/758d3053051511cde2f28fc6b2181b8e227f8ea2.

Torres-Díaz, J. C., Infante Moro, A., & Valdiviezo Díaz, P. (2014). Los MOOC y la masificación personalizada. *Profesorado*, *18*(1), 63–72 http://www.redalyc.org/articulo.oa?id=56730662005.

Umarani, S. D., Raviram, P., & Wahidabanu, R. S. D. (2011). Speech based question recognition of interactive ubiquitous teaching robot using supervised classifier. *International Journal of Engineering and Technology*, *3*(3), 239–243 http://www.enggjournals.com/ijet/docs/IJET11-03-03-35.pdf.

Umer, R., Susnjak, T., Mathrani, A., & Suriadi, S. (2017). On predicting academic performance with process mining in learning analytics. *Journal of Research in Innovative Teaching*, *10*(2), 160–176. https://doi.org/10.1108/JRIT-09-2017-0022.

Vlugter, P., Knott, A., McDonald, J., & Hall, C. (2009). Dialogue-based CALL: A case study on teaching pronouns. *Computer Assisted Language Learning*, *22*(2), 115–131. https://doi.org/10.1080/09588220902778260.

Walsh, K., Tamjidul, H., & Williams, K. (2017). Human machine learning symbiosis. *Journal of Learning in Higher Education*, *13*(1), 55–62 http://cs.uno.edu/~tamjid/pub/2017/JLHE.pdf.

Welham, D. (2008). AI in training (1980–2000): Foundation for the future or misplaced optimism? *British Journal of Educational Technology*, *39*(2), 287–303. https://doi.org/10.1111/j.1467-8535.2008.00818.x.

Weston-Sementelli, J. L., Allen, L. K., & McNamara, D. S. (2018). Comprehension and writing strategy training improves performance on content-specific source-based writing tasks. *International Journal of Artificial Intelligence in Education*, *28*(1), 106–137. https://doi.org/10.1007/s40593-016-0127-7.

Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*, (1st ed.,). Sebastopol: O'Reilly.

Yang, F., Wang, M., Shen, R., & Han, P. (2007). Community-organizing agent: An artificial intelligent system for building learning communities among large numbers of learners. *Computers & Education*, *49*(2), 131–147. https://doi.org/10.1016/j.compedu.2005.04.019.

Yang, Y. F., Wong, W. K., & Yeh, H. C. (2009). Investigating readers' mental maps of references in an online system. *Computers and Education*, *53*(3), 799–808. https://doi.org/10.1016/j.compedu.2009.04.016.

Yoo, J., & Kim, J. (2014). Can online discussion participation predict group project performance? Investigating the roles of linguistic features and participation patterns. *International Journal of Artificial Intelligence in Education*, *24*(1), 8–32 https://www.learntechlib.org/p/155243/.

Yuanyuan, J., & Yajuan, L. (2014). Development of an intelligent teaching system based on 3D technology in the course of digital animation production. *International Journal of Emerging Technologies in Learning*, *9*(9), 81–86. https://doi.org/10.3991/ijet.v11i09.6116.

Zhu, W., Marquez, A., & Yoo, J. (2015). "Engineering economics jeopardy!" Mobile app for university students. *Engineering Economist*, *60*(4), 291–306. https://doi.org/10.1080/0013791X.2015.1067343.

Publisher's Note