Form@re Open Journal per la formazione in rete

articoli

Potential risks of Artificial Intelligence in education

Potenziali rischi dell'Intelligenza Artificiale nell'educazione

Margot Zanetti[a], Stefano Rendina[b], Luigi Piceci[c], Francesco Peluso Cassese[d,1]

[a] *Università degli Studi Niccolò Cusano*, **margot.zanetti@unicusano.it**
[b] *Università degli Studi Niccolò Cusano*, **stefano.rendina@unicusano.it**
[c] *Università degli Studi Niccolò Cusano*, **luigi.piceci@unicusano.it**
[d] *Università degli Studi Niccolò Cusano*, **francesco.peluso@unicusano.it**

Abstract

This work analyses the use of Artificial Intelligence (AI) in education from an interdisciplinary point of view and takes into consideration different potential risks. New studies demonstrated that an AI can *deviate* and become potentially malicious due to various reasons, such as programmers' biases, corrupted feeds, or purposeful actions therefore it seemed necessary to investigate when and how an AI could deviate in educational environment. The first question to pose is if the risks that an AI presents can be applied also to educative forms of AI. The use (or abuse) of the huge amount of sensitive data that these technologies incorporate is one of the most concerning issues. After a review of the increasing literature that deals with the use of technology in classroom, it can be asserted that there exist certain lacks in this research field, and, eventually, authors formulate concrete questions and suggestions to bridge conceptual gaps.

Keywords: technology; artificial intelligence; education; risks; learning.

Sintesi

Questo lavoro analizza l'uso dell'Intelligenza Artificiale (AI) nell'istruzione da un punto di vista interdisciplinare e prende in considerazione diversi potenziali rischi. Nuovi studi hanno dimostrato che un'intelligenza artificiale può *deviare* e diventare potenzialmente dannosa per vari motivi, come i pregiudizi dei programmatori, i feed corrotti o le azioni intenzionali, pertanto è sembrato necessario indagare quando e come un'AI potrebbe deviare anche in ambito educativo. La prima domanda da porsi è se i rischi che presenta un'AI possono essere applicati anche a forme educative di AI. L'uso (o l'abuso) dell'enorme quantità di dati sensibili che queste tecnologie incorporano è una delle questioni più preoccupanti. Dopo una revisione della crescente letteratura che si occupa dell'uso della tecnologia in classe, si può affermare che esistono alcune carenze in questo campo di ricerca e, infine, gli autori formulano domande e suggerimenti concreti per colmare le lacune concettuali.

Parole chiave: tecnologia; intelligenza artificiale; educazione; rischi; apprendimento.

[1] Margot Zanetti is the author of paragraphs 1, 2, 4, 5. Stefano Rendina is the author of paragraph.3. Luigi Piceci: tech consultant and Francesco Peluso Cassese: supervisor.

1. Introduction: education facing the 3th and the 4th Industrial Revolution

The term *industrial revolution* is used to refer to a transformation of productive and social structures brought about by the discovery or innovation of new technologies. The Third Industrial Revolution was determined by the emergence of the internet and the web: the computerisation and interconnection caused by the world wide web and the possibility of working digitally are now showing their powerful effects on society, politics, economy, and education (Penprase, 2018). Those evolutions introduced new classifications, on a sociological level, and put the accent on new questions about technologies and their possible effect on people (Bonfiglio & Piceci, 2019).

From a more technical point of view, universities using telematics as their main means of learning have made large use of the well-known e-learning, u-learning, b-learning, and m-learning (Bidarra & Rusman, 2017). Computer-human interaction studies have developed platforms that can verify the physical and emotional reactions of students, so as to understand their real level of attention and efficiency (Schneider et al., 2016). Moreover, it is well known that emotion is fundamental for an effective and lasting learning (Tyng, Amin, Saad, & Maliket, 2017) and it was wondered what kind of emotional feedback could be given when learning from a screen.

The implications of what is now called the *Fourth Industrial Revolution* seem to be able to make up for this type of shortcoming, as it will be analysed below. This new revolution is determined by the birth and development of new technologies such as biotechnology, nanomaterials, and Artificial Intelligence (AI from now on). AI, understood as the ability of a computer or software to reproduce skills considered typical of the human being, is becoming more and more pervasive into everyday life. This is based on different technologies, such as machine learning, deep learning, neural networking and natural language processing (McDowell Marinchak, Forrest, & Hoanca, 2018). Like any software, it is based on algorithms, which we could simplify as a list of instructions used to solve a problem. There are many attempts to categorize and define AI but, perhaps, the most straightforward is the one proposed by Kaplan and Haenlain (2018): "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (p. 17). Digital-based learning methods are beginning to change by incorporating AI and showing a fascinating process: the machine that must emulate human reasoning, learns from it, and, in turn, it teaches to learn. At the moment, the existing technologies make use of a Weak or Narrow Artificial Intelligence, a kind of intelligence that executes precise instructions. Despite the fact that we are working to reach a Strong Artificial Intelligence, which is one that instead would be able to reproduce in everything and for all human reasoning, including emotions, it seems that the goal is still far away (Alpcan, Erfani, & Leckie, 2017).

This is the reason why this work, which is a conceptual exploration, approached through an argumentation based on selected literature, considers the changes and the challenges brought by the Narrow Artificial Intelligence, now the only one widely used.

2. AI in education: Intelligent tutoring Systems (ITS), content creation and collaborative learning

AI in education is the ability of a system to interpret data entered by teachers and students, to learn from that data, and then to achieve specific learning objectives, constantly adapting. The matters around which the experts of the fields, from a pedagogical perspective, have focused are many: drop-out, self-regulated learning, learning design as educators' professional skills. But the two trends for the contemporary education have been collaborative learning and personalized learning.

The so-called *intelligent support for collaborative learning* is one of the first technologies implemented by AI. Collaborative learning aims to help students to communicate and work with each other. It can be useful to create study groups based on their profiles. There is also the possibility of creating virtual agents, some *facilitators* which can help young people online by simulating another student (Zheng, Niiya, & Warschauer, 2015). An example of this type is Brainly (https://brainly.com/), a knowledge-sharing community where students and experts put their heads together to crack the most difficult homework questions. It is a social learning network in which students can learn from each other also through elements of gamification, in the form of motivational points and ranks.

But this kind of technology has apparently given the possibility to face one of the most recurrent issues of pedagogy: personal tutoring (Kulik & Fletcher, 2016). It is well known that the quality of learning would be better if there were a teacher for each pupil, and it is also known that this is not possible today, because of the costs that would require and due to the structure of the education system itself (Bloom, 1984). This is the main reason why we focus on this kind of technology, together with the fact that is the one that could be considered more subject to risk, as it will be possible to read in the next paragraphs.

In the 1970s and 1980s, a new computer tutoring based on near-born AI and cognitivist theories system began to appear. These systems implemented guided the learner step by step to the solution of a problem, giving suggestions and feedback according to the individual needs, gleaning from a precise database of information. This type of technology is the one that is exploding at the dawn of the Fourth Industrial Revolution and is called Intelligent Tutoring Systems (ITS) (VanLehn, 2011). The ITSs are based on machine learning, namely, on the ability of the machine to improve the effectiveness of algorithms in identifying data patterns based on the number of the entering feeds and on neural network. To date, several private companies provide packages that include ITS, mainly focused on learning the STEM (Science, Technology, Engineering, Mathematics) disciplines, in particular mathematics. The Carnegie Learning Company stands out, claiming that its tools are able to adapt to the needs of each individual student, to make complex pedagogical decisions, and to create a personalized learning path. In this way, students with difficulties can be followed in their problems while those more talented can be stimulated and not slowed down.

These new interactive platforms also amplify the potential of studying the reactions of learners: some of them, as a matter of fact, via webcam, can record the facial expressions of those who are using them to *understand* the emotional reactions. An example of a European matrix is the MaTHiSIS project (http://mathisis-project.eu/). Born to compensate for the difficulties of pupils with special needs, it involves the use of tablets and devices, such as the robot called Nao, which *observes* and *listens* to the student, giving output accordingly.

There are also two other technologies based on AI that are spreading in the world of education: content creation and the augmented and interactive virtual reality. The content creation, offered, for example, by Content Technology Inc., offers the possibility to break down contents of textbooks and to make them easier, to administer self-assessment tests and summaries. Teachers can compose the curricula they want and insert additional content, such as images, videos, and tests.

With augmented virtual reality, it is possible to create learning environments with which it is possible to interact, and, moreover, it is easily usable also for the teaching of humanistic subjects, such as history (Baierle & Gluz, 2018).

Clearly, these technologies require the input of a huge and continuous amount of data. The use of these data has already been questioned (just think about the great research focused on learning analytics; Maseleno et al., 2018), but it has been inquired less on their protection. The AI may be the target of a criminal action and it would be desirable to read something that would deny this hypothesis given the sensitivity of the data inserted in the machine and therefore in the possession of the companies that produce them.

Although this is not the site for a real risk assessment work, some of the risks that any kind of AI can incur will be presented. We will focus on those who can most touch the AI used in education: on the one hand, we will present the ways in which it could more easily *deviate*, namely through different types of bias, and, on the other hand, we will mention the risks related to the amount of data entered.

3. AI: models, datamining and risk from bias

The technologies and the presence of Big Data, based also on data which are involved in areas like those of learning, allow the development of the so-called Learning Analytics, understood as the measurement, the collection, the analysis and the presentation of the data on students for the purpose of understanding and optimizing learning processes and the environments in which this takes place (Ferguson, 2014), linking them to computational aspects (Merceron, Blikstein, & Siemens, 2015).

Leaving aside Academic Analytics, which is not the focus of this article, Ferguson associates Learning Analytics with Educational Data Mining, differentiating them in what these areas of research aim to study: the first focuses on how to improve training, the second is centred on how to extract value from learning data sets. Datasets which today derive mainly from LMS (Learning Management System) platforms, that already have the intrinsic purpose of tracking the student's online learning, as well as from the many existing MOOC (Massive Online Open Courses) courses that share the foundation with the platforms mentioned above.

The collection of data about the student's learning allows to direct the research also on specific cases of learners and to understand how they learn and, through a detailed analysis, to produce models that bring within them characteristics such as knowledge, motivation, meta-cognition, and attitudes of the student.

Educational Data Mining includes the development of e-learning systems (Lara et al., 2014), pedagogical support systems (Hung & Crooks, 2009), grouping of educational data (Chakraborty et al., 2016), student performance forecasts (Kabra & Bichkar, 2011) also concerning deviances (Liu & Hsu, 2013). It is easy to understand how all these data, which are becoming more and more numerous, can generate a sort of profiling of the markers of any kind of student, also allowing a possible realization of training tools based on the involvement of AI. Just think how AI is starting to be used to verify the student's attention, emotions, and conversation dynamics involved in the learning process, as well as for the development and management of the course, such as the attempt to generate optimal groups for collaborative learning activities and foreseeing the dispersion of the students themselves (Nkambou, Azevedo, & Vassileva, 2018). A limited set of data used in the models can, for example, unintentionally perpetuate prejudices existing within our society (Gebru et al., 2019): it should not be forgotten that the algorithm, at least in the early stages, is generated by women/men, anyway, and this may involve a transference not always purified by personal bias.

Biases in AI, even though they are presented in different studies proposed by AI Now Institute, Google and Microsoft (Chou, Murillo, & Ibars, 2017), with different nomenclatures, can be grouped into five categories, or sources, based on their type of distortion:

- dataset bias: the training algorithm for automatic learning is based mainly on the data set that is provided in input, so it is easy to see that if the set is limited and generalized, it can create distortions;
- associations bias: these are the distortions that produce a reinforcement and an amplification of a given prejudice which is present in the dataset. The data provided to the algorithms, by association and analogy, feed the AI and unconsciously train it to continuous prejudices or stereotypes;
- automation bias: they are mainly linked to predictive algorithms and are generated mostly by automatic decisions that the AI itself takes in the phase of self-powering itself;
- interaction bias: voluntary human interaction with AI, after the creation of the algorithm, which can create in the latter a totally distorted learning;
- confirmation bias: intuitively, exactly as it happens in cognitive bias, a confirmation of a bias leads to distorted assumptions and to a strengthening of a generalized belief or distortion, strengthening non-diversified visions.

At the present time, the main studies focus on defining a clear standard for the creation of data from which the algorithms of IA can be generated and fed (Gebru et al., 2019), so as to minimize the presence of bias in the creation phase.

However, what it has been described above thus far, namely Data Mining, Big Data, and Learning Analytics (LA), also falls under the technological governance of Security, Ethics, and Privacy. LA itself collects both personal and sensitive data of students who use learning tools, data on which it is also possible to make profiling in order to provide a specific training program based on the type of learning.

In 2014, from these principles of governance, the European Learning Analytics Community Exchange (LACE), began to point the way for management and the collection about all the data needed to LA, in the protection of Privacy, Security, and Ethics, creating a checklist by the acronym Delicate (Ferguson, Hoel, Scheffel, & Drachslerand, 2016), containing eight action points (1. Determination, 2. Explain, 3. Legitimate, 4. Involve, 5. Consent, 6. Anonymise, 7. Technical aspects, 8. External partners) that must be taken into account by anyone, managers and/or supervisors, who decides to plan the implementation of Learning Analytics solutions (Drachsler & Greller, 2016).

However, in chronological order, the current arrival of the General Data Protection Regulation (GDPR) also extends to the areas of Learning Analytics and Educational Data Mining, with the same principles that are applied to other areas. There is, therefore, a guideline to be followed for everything that concerns the protection, accessibility, and indication of the use that is made of all the data collected.

4. Possible risks of using a *deviate* educational AI: discussion

Therefore, an AI, can *deviate* and this can happen in different ways. Outside the educational field, there have also been several cases that have caused a stir: there has been talk of Google being *racist*, a *sexist* Amazon, and tools to assess the possibility of recidivism prejudicial (Cheng, 2015; Flores et al., 2016). Similar episodes have also happened with regard to image recognition.

Since the relationship between humans and instruments is bidirectional (Vygotskij, 1934/2007), what could the consequences of a possible deviation of the AI used in a classroom be therefore? ITSs, for example, seem promising and they also are welcomed by the school and academic world, but they, nevertheless, present issues that need to be faced. First of all, they are not always efficient (VanLehn, 2011) and one of their primary aims, namely to be able to provide personalized teaching to everyone, is hindered by the different economic and technological development of different states. Nonetheless, more important for this work, the differences are also expressed at a cultural level. As a matter of fact, the language of origin and the culture of belonging influence both the programming of the IA and, logically, that of the feeds that will be introduced by users (Nye, 2014). As it has been presented, these are the primary biases in which the AI can incur, in other words, the causes for which it can deviate. Also, the content creation can suffer from the difference in programming and teaching that each culture could produce, with possible biases.

It must be said, however, that precisely the fact that these technologies are provided mainly by private bodies, it has not been possible to get possession of the algorithms used for programming and therefore it is not known if and to what extent the proposed means are liable to such malfunctions. For the same reason, it is not possible to know the ways in which the data are protected, except on the basis of the general laws on privacy. The relevant questions to education and teaching that seem to be open are therefore the following:

- Can the use of an AI affected by bias, and therefore deviated, transmit the same bias to those who use it?
- The human brain has been forged by millennia of evolution to learn in a certain way. The use of technologies is undoubtedly changing the development of children and the problems related to this are already widely under observation (Cannoni, Scalisi, & Giangrandi, 2018). Learning in this way, skipping a few steps for example when using virtual reality, which shows without really giving true experiencing, what long-term consequences can it have (Barr, 2013)? And, also: what kind of impact can have the interaction with the machine which includes the emotionality on the user?
- Is the huge amount of the constantly entering data, in this case especially belonging to minors, really protected?

For what concerns the possible transmission of bias from machine to human being, different hypotheses can be formulated that see this as more than concrete. First of all, the machine, if affected by bias caused by the data used as training, amplifies the bias of human beings (Levendowski, 2018). The machine, as a matter of fact, which has been trained to reach a certain goal without asking itself moral questions or questions about the best paths to get there, can give answers that most human beings would not give. People, as a matter of fact, are often aware that explicitly expressing their prejudice can be morally wrong, they are ashamed of it and, therefore, they blunt the attitude they would have instinctively had (Zanetti, 2018). The fact that biases are amplified can therefore influence who comes into contact with these biases. The correlation with what recent studies have also highlighted on the transmission of prejudices between people is not considered entirely arbitrary: prejudices with a negative value are exacerbated during the transmission of information (Bebbington, MacLeod, Ellison, & Fay, 2017). It seems that human beings are more receptive to remembering and embracing negative associations. It makes sense, therefore, to believe that a boy who in Google's autofill option, by typing the word *black*, automatically sees offensive suggestions appear may be influenced by it. In a different way, the bias of the machine will also influence female candidates who have seen their curricula automatically discarded by Amazon as women (Zou & Schiebinger, 2018) or the Asian boy who saw his passport not renewed because the applied AI interpreted his photo as if it were with his eyes closed (Cheng, 2016).

There is no reason, then, to doubt that the technologies provided with an AI used in classrooms can present similar risks because project's errors or the not foreseen biases are the same as for the other technologies. Moreover, voluntary actions, aimed at creating confusion or spreading fake news, can be hardly predictable. Unfortunately, to solve these possible malfunctions, that come upstream the use of the devices in schools, there is nothing that teachers and students could do. However, it could be suggested to technological corporations to form interdisciplinary teams that include also the teachers for the planning and monitoring phases.

The issue of the influence on natural brain development on long term would require longitudinal studies which there does not seem to be any trace of at the moment. What is interesting in this regard is the possible impact that these intelligences can have on an emotional level. The Global Risk Report (GGR) 2019 (World Economic Forum, 2019), as a matter of fact, individuates as one of the major risk factors related to the use of these technologies exactly the potential effects that an emotional interaction with them can lead to. Specifically, the GRR analyses in detail the dangers of radicalization, dependence, decreased ability to control impulses, lack of relational skills, and a greater predisposition to various pathologies, which are nothing more than the exacerbations of what is already presented in the literature. AI cannot understand the context in which a sentence is said, the metaphors, or facial expressions related to an emotion, but it can learn them, as new researches are showing. The interaction between an authentic emotionality and a *fictitious* one seems to be a risk. As written above, the recent platforms using AI aim precisely at perfecting the recognition of the learner's emotions and at an increasingly sophisticated interaction to make him/her learn (see human-robot Interaction; e.g. Erol et al., 2019).

This should not lead us to believe that the use of these technologies is opposed or highly dangerous, even if there is no real risk assessment research. Privacy regulations, ethical codes and methods for reducing biases are among the main subjects of study and research. Of course, given the possible complications that these means can bring, we cannot but hope that the schools that intend to use them will promote specific training courses for teachers, so that they can better supervise the activities and integrate them with others designed to make up for the scientifically proven shortcomings.

5. Conclusion: emergent pedagogical necessities

Recent research has shown that AI can deviate and that can do so on the basis of what it is fed. Nevertheless, it would be foolish not to take advantage of the means that seem to be able to make up for shortcomings which have already been identified.

At the dawn of the Fourth Industrial Revolution, providing a personalized teacher to each student with ITS, creating and organizing textbook content, immersing oneself in an interactive virtual reality and finding ways to make the students collaborate more, seems to be something highly desirable. But if and when doing so, it is necessary to consider all the possible consequences. And here it is one of the first emergencies not only of pedagogy, but of all disciplines: dialogue and collaboration.

Different types of biases in which the machine can incur have been exposed and moreover it has been raised issues related to the protection of data, which are used today mainly in Learning Analytics. This technical information was then applied to everyday reality, especially in relation to the Artificial Intelligence in Education (AIEd), and this has led to the emergency, in the writer's opinion, of problems that urge to be taken into account and to new challenges for pedagogy. If biases are transmissible and the interaction with IA can have such a great impact on the training and development of young people, it seems necessary to implement ethical and critical thinking and give the tools to self-form in a society that requires continuous training, especially with regard to technologies (Penprase, 2018). First of all, it can be extremely useful if the teacher who uses these services is well prepared on their functioning and possible risks so that they can do their best to prevent or recognize them, if necessary. It is important to monitor the effects on pupils and to support educational pathways aimed at reducing the possibility of concretisation of such risks. According to some authors, today's teaching can be defined as *Anthropocentric Humanism* and it is hoped that in order to face the emerging realities it will become "critical posthumanism" (Jandric, 2017). By this we mean that the division and gap between technical or STEM subjects and the classical ones must be filled again: every knowledge is necessary to the other and to the increasing technical knowledge; for a real progress, it is necessary to associate a stronger and stronger humanistic knowledge. In this regard, we point out an interesting reality launched by Stanford University, the Human-Centered Artificial Intelligence Institute (https://hai.stanford.edu/), which aims to fill many of the gaps that have been highlighted in this article through interdisciplinary work and its purpose is to investigate the impact that the IA has on humans.

Reference list

Alpcan, T., Erfani, S. M., & Leckie, C. (2017). Toward the starting line: A systems engineering approach to strong AI. https://arxiv.org/abs/1707.09095 (ver. 23.03.2020).

Baierle, I. L. F., & Gluz, J. C. (2018). Programming intelligent embodied pedagogical agents to teach the beginning of the industrial revolution. In R. Nkambou, R. Azevedo, & J. Vassileva, (Eds.), *Intelligent Tutoring Systems. ITS 2018* (pp. 3-12). Cham: Lecture Notes in Computer Science, Springer.

Barr, R. (2013). Memory constraints on infant learning from picture books, television on touchscreens. *Child Development Perspectives*, *7*(4), 105–110.

Bebbington, K., MacLeod, C., Ellison, M., & Fay, N. (2017). The sky is falling: Evidence of negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92–101.

Bidarra, J., & Rusman, E. (2017). Towards a pedagogical model for science education: bridging educational contexts through a blended learning approach. *Open*

Learning: The Journal of Open, Distance and e-Learning, *32*(1), 6–20. https://doi.org/10.1080/02680513.2016.1265442 (ver. 23.03.2020).

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*(6), 4–16.

Bonfiglio, L., & Piceci, L. (2019). Interventi precoci sulle difficoltà dell'apprendimento attraverso una comunicazione adatta alle generazioni digitali: didattica capovolta e tecnologie. *Giornale Italiano di Educazione alla Salute, Sport e Didattica Inclusiva*, *1*(3), 28–45. https://doi.org/10.32043/gsd.v1i1.107 (ver. 23.03.2020).

Brainly. https://brainly.com/ (ver. 23.03.2020).

Cannoni, E., Scalisi, T. G., & Giangrande, A. (2018). Indagine sui bambini di 5-6 anni che usano quotidianamente i dispositivi mobili in ambito familiare: caratteristiche personali e contestuali e problematiche cognitive ed emotive. *Rassegna Italiana di Psicologia*, *35*(1), 41–56.

Chakraborty, B., Chakma, K., & Mukherjee, A. (2016, March). A density-based clustering algorithm and experiments on student dataset with noises using Rough set theory. In *2016 IEEE International Conference on Engineering and Technology (ICETECH)*, pp. 431-436). IEEE.

Cheng, S. (7 dicembre 2016). An algorithm rejected an Asian man's passport photo for having 'closed eyes'. *Quartz*. https://qz.com/857122/an-algorithm-rejected-an-asian-mans-passport-photo-for-having-closedeyes/ (ver. 23.03.2020).

Chou, J., Murillo, O., Ibars, R. (26 settembre 2017). How to Recognize Exclusion in AI. *Medium*. https://medium.com/microsoft-design/how-to-recognize-exclusion-in-ai-ec2d6d89f850 (ver. 23.03.2020).

Drachsler, H., & Greller, W. (2016). Privacy and analytics – it's a DELICATE issue. A checklist to establish trusted learning analytics. *6th Learning Analytics and Knowledge Conference 2016*, Edinburgh, UK.

Erol, B. A., Majumdar, A., Benavidez, P., Rad, P., Choo, K. R., & Jamshidi, M. (2019). Toward Artificial Emotional Intelligence for Cooperative Social Human–Machine Interaction. *IEEE Transactions on Computational Social Systems*, *7*(1), 234–246.

Ferguson, R. (2014). Learning analytics: Fattori trainanti, sviluppi e sfide. *TD tecnologie didattiche*, *22*(3), 138–147.

Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of Learning Analytics*, *3*(1), 5–15. https://doi.org/10.18608/jla.2016.31.2 (ver. 23.03.2020).

Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False Positives, False Negatives, and False Analyses: A Rejoinder to Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks. *Federal Probation*, *80*, 38–46.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., & Crawford, K. (2019). *Datasheets for Datasets*. https://arxiv.org/abs/1803.09010 (ver. 23.03.2020).

Human-Centered Artificial Intelligence Institute. Stanford University. https://hai.stanford.edu/ (ver. 23.03.2020).

Hung, J. L., & Crooks, S. M. (2009). Examining online learning patterns with data mining techniques in peer-moderated and teacher-moderated courses. *Journal of Educational Computing Research, 40*(2), 183–210.

Jandric, P. (2017). From Anthropocentric Humanism to Critical Posthumanism in Digital Education. In P. Jandrić, *Learning in the Age of Digital Reason*, pp. 195-210. Rotterdam: Sense Publishers.

Kabra, R. R., & Bichkar, R. S. (2011). Performance prediction of engineering students using decision trees. *International Journal of computer applications*, *36*(11), 8–12.

Kaplan, A., & Haenlein, M. (2018). 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.

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A meta-analytical review. *Review of Educational Research*, *86*(1), 42–78.

Lara, J. A., Lizcano, D., Martínez, M. A., Pazos, J., & Riera, T. (2014). A system for knowledge discovery in e-learning environments within the European Higher Education Area–Application to student data from Open University of Madrid, UDIMA. *Computers & Education*, *72*, 23–36.

Levendowski, A. (2018). How copyright law can fix artificial intelligence's implicit bias problem. *Washington Law Review*, *93*(2), 579–630.

Liu, Y., & Hsu, Y. (2013). Predicting adolescent deviant behaviors through data mining techniques. *Journal of Educational Technology & Society*, *16*(1), 295–308.

Maseleno, A., Sabani, N., Huda, M., Ahmad, R., Jasmi, K. A., & Basiron, B. (2018). Demystifying learning analytics in personalised learning. *International Journal of Engineering & Technology*, *7*(3), 1124–1129. https://doi.org/10.14419/ijet.v7i3.9789 (ver. 23.03.2020).

MaTHiSIS project. http://mathisis-project.eu/ (ver. 23.03.2020).

Marinchak, C. L. M., Forrest, E., & Hoanca, B. (2018). The impact of artificial intelligence and virtual personal assistants on marketing. In M. Khosrow-Pour, DBA (Ed.), *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 5748-5756). Hershey, PA: IGI global.

Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning analytics: From big data to meaningful data. *Journal of Learning Analytics*, *2*(3), 4–8.

Nkambou, R., Azevedo, R., & Vassileva, J. (2018). Intelligent Tutoring Systems. *14th International Conference, ITS 2018*, Montreal, QC, Canada. https://link.springer.com/book/10.1007/978-3-319-91464-0 (ver. 23.03.2020).

Nye, B. D. (2014). Intelligent Tutoring Systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, *25*(2), 177–203.

Penprase, B. E. (2018). The fourth industrial revolution and higher education. In N. W. Gleason (Ed.), *Higher education in the era of the fourth industrial revolution* (pp. 207-229). Yale-NUS College, Singapore: Palgrave McMillan.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2016). Using mobile eye-trackers to unpack the perceptual benefits of a tangible user interface for collaborative learning. *ACM Transaction on Computer-Human Interaction*, *23*(6), 1–23. https://doi.org/10.1145/3012009 (ver. 23.03.2020).

Tyng, C. M., Amin, H. U., Saad, M. N., Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.01454 (ver. 23.03.2020).

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369 (ver. 23.03.2020).

Vygotskij, L. S. (2007). *Pensiero e linguaggio*. Firenze: Giunti. (Original work published 1934).

World Economic Forum (2019). *The Global Risk Report 2019*. https://www.weforum.org/reports/the-global-risks-report-2019 (ver. 23.03.2020).

Zanetti, M. (2018). Pregiudizio ed etichettamento: il ruolo dell'insegnante nello sviluppo di comportamenti devianti. *Formazione & Insegnamento*, *16*(2), 193–204.

Zheng, B., Niiya, M., & Warschauer, M. (2015). Wikis and collaborative learning in higher education. Technology. *Pedagogy and Education*, *24*(3), 357–374.

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist - it's time to make it fair. Nature, 559, 324–326.