SCIENTIFIC PUBLICATIONS OF THE STATE UNIVERSITY OF NOVI PAZAR
SER. A: APPL. MATH. INFORM. AND MECH. vol. 15, 2, 2023, 73–85

Ethical Challenges in Open Learning Analytics

I. Fetahović, E. Mekić, K. Kuk, B. Popović, E. Dolićanin

Abstract: Open learning analytics (OLA) is a new research field focusing to create an open platform for integration of heterogeneous learning environments. The central notion in OLA is openness, and it relates to architectures, processes, access, and datasets. Learning analytics (LA) usually employs artificial intelligence (AI) technologies and machine learning (ML) algorithms to provide information and predictions, thus enhancing learning experience. AI-based systems provide many benefits, but their implementation raises ethical concerns due to wide range of possible negative consequences. OLA approach amplifies ethical concerns related to AI-based systems, such as privacy, transparency, bias, fairness, etc. In this paper we discuss a comprehensive list of ethical issues in OLA along with mitigations procedures.

Keywords: open learning analytics, ethics, artificial intelligence, machine learning, education

1 Introduction

The rapid advancement in AI has facilitated the implementation of AI in education (AIED) applications. AIED refers to the use of AI technologies in educational settings to enhance teaching, learning, or decision making, by providing valuable information and predictions to all stakeholders. Students receive personalized guidance and support, while teachers and policymakers use AI to help them in making decisions. Students are increasingly using AI and ML through various applications such as ones that provide enhanced learning through online learning platforms, personalized education environments or LA. The new generation of online learning platforms such as Learning Management Systems (LMS) or Massive Open Online Course (MOOC) leverage AI technologies to achieve personalized learning.

AI and ML applications are incredibly influential, but also have a potential to cause harm. Researchers have shown that prominent ML algorithms, often considered neutral, can be biased against women, ethnic minorities, and low-income individuals [1] .

LA is an application of data analytics in education. It creates and uses procedures and tools to collect, analyze and report the data about learners and their contexts, in order to support and enhance the learning process. LA aims to monitor and analyze educational

Manuscript received Jul 14, 2023; accepted October 11, 2023.

I. Fetahović, E. Mekić and E. Dolićanin are with State University of Novi Pazar, Novi Pazar, Serbia; K. Kuk and B. Popović are with University of criminal investigation and police studies, Belgrade, Serbia

data to provide assessment and feedback, predict certain outcomes, or create recommendations and personalized educational experience. However, LA may also unintentionally result in consequences that benefit some students while harming others. OLA takes LA one step further. It intends to integrate heterogeneous learning environments through concept of openness, which refers to protocols, architectures, software, algorithms, or datasets.

In this paper, we analyze and report ethical issues and challenges in LA associated with different learning environments and corresponding technologies. We also describe ethical problems that arise when connecting different data sources, environments, and technologies in the context of OLA. Special attention is given to the ethical issues in ML and AI as underlying technology for LA.

2 Background

Decisions produced by AI systems affect every aspect of our lives. Interest for embedding ML and AI algorithms into digital technologies has grown over the years due to many benefits, such as reduced costs and increased productivity. ML algorithms are used by educational and health institutions, financial sector, and national governments [2]. AI has been applied to various domains, such as pattern and voice recognition [3], decision-making, natural language processing [4], and automatic language translation. Furthermore, individuals often use recommender systems to help them choose a book, song or a product [5].

Ethics is a branch of philosophy concerned with what is right for individuals and society [6]. It can be defined as a moral code of norms that exist in society externally to a person, depending on culture and time. AI ethics deals with moral, political and social implications of design, implementation and use of AI based systems. AI may cause unintended negative consequences, abuses and wide range of both individual and societal harms.

The ethics of AI in general has received a great deal of attention in the research community, political institutions and civil societies [7]. Many of these research efforts and initiatives focus on data related issues - informed consent, data privacy, biased datasets and biased assumptions, and transparency.

Jobin et al. [8] analyzed the corpus of 84 documents containing principles and guidelines on ethical AI and reported a global convergence around five ethical principles - transparency, justice and fairness, non-maleficence, responsibility, and privacy, with substantive divergence to how these principles are interpreted. The most frequently mentioned ethical issue in the literature is privacy and data protection. AI often needs large datasets for training and accessing those data can raise concerns of data protection. Studies have shown that AI may pose privacy risks even where no direct access to data is achieved [9]. AI has the potential for identification of anonymized personal data and can create new data protection risks not envisaged by legislation [10].

Munn [11] argues that ethical principles and guidelines for AI technologies that emerged in recent years are meaningless, incoherent, and eventually difficult to apply. He advocates for alternative and intersectional approaches which require considering a wide array of social and political factors.

3 Open Learning Analytics

The term open learning refers to the fact that all or most of the teaching is done by someone who is separated in time and space from the learner, whose learning and supporting material is available to anyone, as well as the mission's desire to include more dimensions of openness and flexibility. The two most important aspects of openness are free availability over the Internet and as few restrictions on resource use as possible, whether technical, legal, or financial. This concept gives the student control over his or her learning path and style. Openness is not associated only with learning material but is connected with all aspects of learning (management, teaching, marking, learning results analysis, etc.) [12].

LA is a well-known field focused on the collection, measurement, analysis and reporting of educational data to optimize learning and the environment in which it occurs, thus providing benefits for students, instructors, and institutions. LA may be applied to support students through self-assessments, recommender systems, visualizations, personalized learning paths, and real-time feedback, but also to support institutions in resource allocation, student success, and finance [13]. LA is an interdisciplinary field between data science/AI, statistics, learning sciences, and psychology. Majority of LA approaches so far have been narrow and restricted to specific learning environments. Improving LA means scaling it up to include distributed datasets across a variety of different platforms and environments, to address the needs and objectives of various stakeholders.

OLA is a relatively new research area that refers to an open platform to integrate heterogeneous LA approaches. The term openness is yet not clearly defined, but it usually refers to many things concerning LA, such as 1) architecture, processes, algorithms, tools, and methods; 2) access and participation; 3) datasets, etc [14]. OLA needs open-source software that uses open procedures, algorithms, and technologies to make it simple for researchers and developers to integrate their own tools and approaches with the platform. OLA can be viewed as a set of tools that will bring together the fields of LA and open source software development as a means to investigate the intersection of LA and open learning, open technologies, and open research [15]. The research community in OLA emphasizes two major needs: 1) the development of appropriate open-source software, open standards, and open APIs to address interoperability issues in this field, and 2) the importance of addressing ethical and privacy issues [16]. OLA uses data originating from multiple sources to address the needs and objectives of different stakeholders by applying suitable analytics tools and procedures.

4 Ethical Issues in OLA

The advancement in applying digital tools and technologies in learning, as well as the availability of educational media, facilitated the emergence of a massive amount of educational datasets from a wide variety of sources and environments. Datasets can originate from centralized educational environments, open and networked learning environments, messaging platforms, social media, and sensory data coming from handheld and wearable devices. All of these learning environments and associated technologies have their own security and ethical issues. However, connecting and integrating them together amplifies many of them, such as transparency and data privacy.

Different types of data can be collected from educational data sources. It can be activity data, assessment data, user profile data, etc. OLA approach implies not only using and storing these data but transferring them between many different environments. However, many of these data types can be classified as sensitive information and should be treated with special attention. Mismanagement and inappropriate handling of data may cause serious harm to stakeholders, especially adolescent learners and vulnerable groups, so OLA solutions must incorporate mechanisms and procedures to avoid these risks. OLA solutions must provide detailed documentation on how the data is being collected, processed, and stored, how the data will be used, and how the learner's identity will be protected. This is called transparency and is crucial for the legitimacy and acceptance of OLA. Stakeholders must also understand the purpose and benefits of using OLA.

LA can be performed in various ways, including basic data visualization and statistical analysis. However, contemporary LA tools and methods employ AI and ML algorithms as underlying technology. Therefore, when discussing ethical issues in LA, it should be done in relation to ethical problems in the field of AI and ML.

Ethics of LA is an emerging field of study and it still remains unclear what it should include. It usually includes ethical issues such as: informed consent and privacy, the interpretation of data, the management of data, as well as on much broader issues such as power relations, surveillance, and the purpose of education [17].

In the remainder of this section we discuss the most important ethical issues in OLA.

Transparency

Transparency is a major issue when discussing the ethics of AI and ML algorithms. It means developing algorithms where all involved parties are aware of their inner workings, assumptions, etc. Transparency includes the ability to explain and interpret model predictions and behaviour. ML systems are inherently not transparent, and the commercial confidentiality of models may further limit transparency. Lack of transparency may be inherent due to technology limitations or acquired by design decision and/or obfuscation of the underlying data [18]. Modern ML algorithms are complex and fed with large amounts of data which makes them almost impossible to read or interpret due to human cognitive limitations or lack of appropriate visualization tools. Lack of transparency can also result from the malleability of algorithms, whereby algorithms can be continuously reprogrammed for improvement, thus obfuscating the history of its evolution [19].

Transparency is a major issue in OLA context. OLA predicts use of open source software, which promotes transparency. However, open standard, protocols, software and datasets contribute to system complexity, opaqueness and hinder explainability.

There are different ways of addressing the problems related to lack of transparency. The use of open-source software in the context of ML transparency has already been advocated for over a decade [20]. Gebru et al. propose using documentation to describe "every component, no matter how simple or complex, is accompanied with a datasheet describing its operating characteristics, test results, recommended usage, and other information" [21]. This might be difficult to implement, especially for proprietary algorithmic systems. Furthermore, there is no standard format for documenting the origin of a dataset [19]. Another approach to transparency is the use of technical tools to test and audit algorithmic systems for negative impacts and auditing a prediction or decision process in detail [22]. Every party in OLA context should commit to strong testing and auditing tools and procedures describe above. However, OLA also needs collaborative auditing, which requires sharing of relevant information on algorithms, datasets and data governance procedures between all involved parties.

Transparency is related to notions of explainability and interpretability. Some ML models, especially the complex ones, such as those relying on deep neural networks, produce results that cannot be easily interpreted by humans. Such opaqueness and lack of transparency of the model may be problematic or even unacceptable to humans, especially to those who are directly affected by the decisions of the model. This issue is further amplified in situations when input data contains traces of bias and inequality. Most models focus on the improvement of their used metrics, while less considering the interpretability. Interpreting and explaining the workings and decisions reached by an algorithmic model in OLA means understanding why those outcomes are produced, and having the ability to present to the affected individuals the rationale behind that decision as if it had been produced by human. We need to be able to understand and explain model's outcomes in terms of its logic, technical rationale, social context, and moral justifications. The first two components of interpretable AI are important in terms of technical consideration, while the last two are significant at the point of delivery and deployment. Defining social context means translating technical jargon of algorithmic decision process and outcomes into language of the humanly relevant categories, relationships and meanings, which relates to the model purpose and objectives.

The central notion for applying transparency is informed consent, which refers to an individual permitting data collection and allowing action to be taken based on algorithmic predictions. Institutions should design appropriate informed consent forms that learners must sign before they can participate in OLA context. Those who decline participation should not be disadvantaged in any way. Institution must inform and convince the potential learners that the OLA project in question is beneficial, innovative and fair, and that advantages outweigh the risk. In the modern era of IoT and smart sensors, collection of private information is possible without individual's awareness and consent. Institutions in OLA environment should identify a person who will receive and handle these complaints.

Accountability, responsibility and assessment

AI systems may take over functions that were previously done exclusively by human agents. In a situation of negative consequences of AI systems' decisions and predictions, it may be difficult to designate responsibility and determine accountable parties, thus leading to the violation of rights of the affected individuals or groups. This is further amplified in OLA context, where different AI-enabled software systems produce and share information between each other. These system are operated by many individuals belong to educational institutions with different procedures for determining responsibility. Moreover, different sites within OLA context may belong to different countries and cultures with various views and notions on responsibility and accountability.

Accountability is a principle of identifying and assigning responsibilities for a proper functioning of every component in OLA systems. Issue of accountability is divided into two components: answerability and auditability. Answerability demands a clear and continuous chain of human responsibility to be constructed within OLA context. Auditability deals with the question of how the people designated as the responsible ones are to be held accountable.

Accountability gap arises from the fact that AI systems can't be held responsible for judgments and decision which affect the lives of others as humans do. This must be addressed in a way to complement AI decisions with a relevant source of human answerability. This is not a simple task in OLA due to complexity and collaborative character and using AI systems in OLA context, so it may be difficult to establish the levels of responsibility in it. OLA necessitates sharing the results of review process with other parties in the OLA context to better address issues of accountability and responsibility. Lack of transparency and explainability may undermine moral responsibility and accountability for the actions performed by ML algorithms in OLA context. Complexity of ML algorithms may lead to denial of responsibility or "agency laundering", where one blames the algorithm for any controversial actions. To address this issue, establishing ethical committee for algorithms' oversight is advocated in [23], but critics stress that it can lead to "ethics bluewashing" – presenting ML system as more ethical than it is [24]. Buhmann et al. [25] suggest that organization should take responsibility for their actions, while Floridi [26] introduces a concept of distributed moral responsibility that assigns moral responsibility to all relevant parties in a network. Establishing appropriate ethical committee consisting of members from all sites within OLA context is necessary for collaborative and distributed approach to responsibility.

Bias

Algorithms are not ethically neutral, and may be biased in various ways, such as gender bias and racial bias. Machine bias may cause harm by creating or reproducing existing discriminations based on gender, sex, or ethnicity, especially when it comes to minority groups. It can harm human life in various ways, ranging from subtle impact to serious discrimination with psychological, financial or life and death consequences.

An unfair or biased algorithm is one whose decisions are skewed toward a particular group of people. It is mostly caused by the prejudices of the model programmers or inappropriate training datasets. Design, structure and features of AI based systems and models for data mining, analytics and prediction are generated by their designers, and may suffer from their designers' preconceived notions, prejudices and biases. Data sources for building and training AI and ML systems correspond to the current social structure and dynamics, and may reinforce and amplify existing patterns of discrimination and inequality. ML model can be fed with preferentially sampled, inacurate, skewed, data or with data reflecting existing systemic societal bias [27]. Datasets for the training of ML algorithms can be imbalanced, so they don't correctly resemble or represent the entire population. This can result in flawed, biased or discriminatory inferences and predictions. OLA systems are prone to bias, partially due to open approach towards datasets and involvement to institutions and people from potentially various countries and cultures with different prejudices. Discrimination on the basis of certain racial, sexual, ethnic or other characteristic is not just an ethical issue but also a violation of human rights laws and regulations, which makes it illegal in many jurisdictions.

The best practical approach to tackle bias is to collect better data that includes sufficient proportions of all groups of interest. Researchers [28] suggest that explicit racial bias in algorithms can be mitigated by existing laws and regulations, while implicit, unconscious bias necessitates workplace diversity within high-tech industries and appropriate public policies for bias detection and mitigation. OLA is especially sensitive to implicit bias, due to difference in culture where involved parties may not be aware of stereotypes. Due to collaborative nature of OLA and sharing of information, this stereotypes and prejudices may propagate to other institutions.

Another approach to mitigating bias is to exclude sensitive data variables containing information about gender or race. The processing of such information is limited or prohibited under anti-discrimination and data protection law, to limit the risks of unfair discrimination. Data control procedures and excluding sensitive data from sharing might be suitable approach for OLA. This is beneficial not only to possible bias propagation, but also to data privacy concern.

Fairness

Fairness is an important notion in AI systems since those systems are built by people who are generally prone to biases and prejudices. In the context of AI, fairness usually refers to mitigating bias, discrimination, and harm. Fairness is an important feature of an algorithm for preventing discrimination. However, the notion is subtle and ambiguous, and researchers can't agree on the definition, measurements and standards of algorithmic fairness. Authors in [29] suggest redefining concept of fairness of ML algorithms to include notion of respect. Three distinct approaches for algorithm fairness are being proposed in the research community: statistical parity, conditional statistical parity and predictive equality [30]. Considering historical and sociological context not present in the ML datasets may be helpful in determining appropriate approaches to fairness in algorithms.

Veale and Binns [31] propose a third-party intervention and assessment for improving algorithmic fairness. External entity holds sensitive data and analyses the data and the model for potential traces of discrimination. We suggest this approach as the most appropriate in the OLA applications. External entity for OLA should also receive auditing information from all educational institutions and conduct algorithm fairness analysis.

Labelling and profiling

Analysis in [32] has shown that profiling can give rise to problems at the interface of law and ethics. Author concluded that profiling is potentially inconsistent with the individual rights of citizens in liberal democracies. LA may use a students' historical data to categorize them based on their predicted degree of success which can limit learning, discourage and manipulate students, and lead to stereotypes, mistreatment and discrimination. Profiling is a situation when ML algorithms categorize people based on characteristics such as location, age, and behavior. Learning environment should allow students to learn from their mistakes and past experiences without being labeled. Institutions in OLA context should create auditing procedures for tracking labeling and profiling and propose measures to mitigate negative consequences of these practices. We recommended these procedures to be uniform for all involved parties in OLA context.

Data ownership, control and sharing

Data ownership is a complicated ethical issue that relates to both raw and analyzed data. Central question is who owns the data—the institutions, the students, or the companies that use the data to create LA tools. Students should be able to decide and force institutions to delete the data that are no longer needed by the system through the right to be forgotten. Types of data to support LA and AI in education include recording of students' outcomes and competencies, their emotional states, and other sensitive data. The question is not only who owns and who is able to access these data, but how the data should be processed, analyzed and shared. Some authors claim that in the coming years it would be impossible to maintain personal privacy and control at scale, so they emphasize ethical usage of the data with ethical guidelines that are clearly understood [33]. OLA includes sharing data between systems and institutions. Keeping track of data transfer and sharing might be cumbersome, so rational data privacy rules and guidelines might be difficult to construct and apply. OLA institutions should provide uniform guidelines for ethical usage of data to be applied at all institutions in OLA context.

Autonomy

ML algorithms may hinder human autonomy and undermine human dignity and self determination. Human autonomy should be encouraged and model's autonomy constrained or reversed. Automation bias happens when human decision maker ignore their own assessment on behalf of ML analytics due to its perceived objectivity. Promoting human autonomy is possible through participation and negotiation of different stakeholders in the design process. Addressing this issue in OLA may be done through disclaimers provider by learning environments and educational institutions to teachers, emphasizing statistical nature of ML algorithms and important of human agency in decision making process.

Privacy

Implementation of AI based solutions often includes the process of capturing and extracting personal data without the proper consent of the individual. Moreover, improper handling of the data poses a risk of revealing such data to unauthorized third parties. Some AI systems have the ability to target and profile individuals without their knowledge and consent which can be interpreted as violating their basic rights to privacy. ML algorithms reduce people's ability to control who has access to information and how the information is further processed and stored. When discussing privacy, some other connected issues naturally arise, such as data protection and security, anonymity and informed consent. Users may not be aware how and what type of information is being collected as they use ML based systems.

Resolving data ownership and data control issues can help to establish good balance between data privacy and data-processing benefits. Regulations such as GDPR have provided important procedures to tackle these problems. Blockchain, decentralized identifiers (DIDs), and Smart Contracts are three Web 3 technologies that can be used to increase privacy. Smart contracts can be used to automate processes in OLA, such as course enrolment, completion, and certification. This technology can also be used to establish trust and transparency in the education system.

Stakeholders addressed privacy through trust, but some stakeholders may not trust one another in OLA. Privacy is also connected with data control and governance. For example, in USA there is no single federal law regulating information privacy, but a system of federal and state regulations that contradicts one another [34]. On the other side, EU launched the General Data Protection Regulation (GDPR) which contains strict standards and rules related to privacy, transparency and informed consent. GDPR applies to HEI outside of EU, if they offer courses to EU citizens. OLA context implies intensive data access and data manipulation activities such as sharing, transforming and processing. This is contradictory to data privacy principles. Moreover, learner's data in OLA is usually stored outside of their home institutions or in a different country with potentially different laws and cultural norms regarding privacy. Due to these complexities, OLA guidelines should focus on ethical usage of data instead on strict data privacy rules.

Social isolation, surveillance

Excessive usage of modern information technologies decreases the need for social interaction and may lead to social isolation and can cause mental problems since humans evolved as social animals that require frequent social interactions and understanding. Personalized solutions enabled by AI algorithms may result in hyper-personalization, when the individual's exposure to different opinions and attitudes is limited. This can negatively influence some human capacities, such as abilities to establish trust and empathy, thus polarizing social relationships and degrading social cohesion, which are crucial for well-ordered societies.

Surveillance in learning environments has a potential to alter behavior of not only students, but also of teachers, faculty and administrators, transforming existing power relations and creating new roles and responsibilities [17].

The obligation to act

Institutions should strive to support and encourage students, while students have a responsibility to do their best to succeed. Institutions are ethically obliged to act in accordance with data provided by OLA, and students should share their learning data with OLA, if the system is reliable and fair.

Conclusion

The proliferation of technology-enabled learning created a context in which educational stakeholders generate massive amounts of data across different learning environments. LA and big data analytics can leverage this context to develop a greater understanding of the learning experience. OLA leverages open approach towards data, algorithms, and protocols to enhance LA.

In this paper we have discussed ethical challenges in OLA. We have provided a comprehensive list of ethical issues and corresponding mitigation procedures. However, ethics of AI in education must address not only data or computation related issues, but also questions of the ethics of education [35]. These questions include, but are not limited to: the purpose of learning, the choice of pedagogy, the relationship between teachers and technology, and access to education.

Future research should be oriented towards establishing guidelines and frameworks for ethical challenges in OLA. Framework should incorporate new knowledge and understandings about learning and teaching, as our science, value, morality change over time [36]. It should consider issues related to data and algorithms, ethics of education, and overlaps between them. Full ethics framework has also to consider issues that have yet to be even identified, such as new insights from pedagogy, psychology, educational neuroscience, and philosophy.

References

[1] J. BUOLAMWINI, T. GEBRU, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, Proceedings of Machine Learning Research, 81, 1–15, 2018.

[2] H. ROBERTS ET AL., *The Chinese approach to artificial intelligence: an analysis of policy and regulation*, AI & society, 36, 59–77, 2021.

[3] O. ESKICIOGLU, E. DOLICANIN, A. ISIK, & R. KUCI, *Recognition and Detection with Deep Learning Method*, Scientific Publications of the State University of Novi Pazar, Series A: Applied Mathematics, Informatics & Mechanics, 13(2), 105-115, 2021.

[4] A. AVDIC, U. MAROVAC, D. JANKOVIC & S. MAROVAC, *Automatic Labeling of Diagnosis in Medical Reports in Serbian*, Scientific Publications of the State University of Novi Pazar, Series A: Applied Mathematics, Informatics & Mechanics, 14(2), 65-72, 2022.

[5] D. PARASCHAKIS, *Towards an ethical recommendation framework*,In: 2017 11th International Conference on Research Challenges in Information Science (RCIS), 211–20, Brighton, United Kingdom, 2017.

[6] C.M. GRAY, E. BOLING, *Inscribing ethics and values in designs for learning: a problematic*, Education Technology Research and Development, 64, 969–1001, 2016.

[7] P. BODDINGTON, *Towards a Code of Ethics for Artificial Intelligence*, Cham: Springer, 2017.

[8] A. JOBIN, M. IENCA, E. VAYENA, *The global landscape of AI ethics guidelines*, Nature machine intelligence, 1(9), 389–399, 2019.

[9] C. JERNIGAN, B.F. MISTREE, *Gaydar: Facebook friendships expose sexual orientation*, First Monday, 14(10), 2009.

[10] J. TAO, T. TAN, *Affective computing: A review*, Proceedings of the First International Conference: Affective computing and intelligent interaction, ACII 2005, Beijing, China, 981-995, 2005.

[11] L. MUNN, *The uselessness of AI ethics*, AI and Ethics, 3(3), 869–877, 2023.

[12] A. BOZKURT, *From distance education to open and distance learning: A holistic evaluation of history, definitions, and theories*, In: Handbook of Research on Learning in the Age of Transhumanism, 252-273, IGI Global, 2019.

[13] P. LEITNER, M. KHALIL. M. EBNER, *Learning Analytics in Higher Education—A Literature Review*, In: Peña-Ayala, A. (eds) Learning Analytics: Fundaments, Applications, and Trends. Studies in Systems, Decision and Control, 94, Springer, Cham, 2017.

[14] A. MUSLIM, M.A. CHATTI, M. GUESMI, *Open Learning Analytics: A Systematic Literature Review and Future Perspectives* , In N. Pickwart & S. Liu (Eds.), Artificial intelligence supported educational technologies: Advances in analytics for learning and teaching, 93–114, Springer, 2020.

[15] S. ALEXANDER ET AL., *OLA Press Release* , Retrieved from http://solaresearch.org/initiatives/ola/ 2014.

[16] M.A. CHATTI, A. MUSLIM, U. SCHROEDER, *Toward an Open Learning Analytics Ecosystem*, In: Kei Daniel, B. (eds) Big Data and Learning Analytics in Higher Education, Springer, Cham, 2017.

[17] S. SLADE, P. PRINSLOO, *Learning analytics: Ethical issues and dilemmas*, America Behavioral Scientist, 57 (10), 1509–1528, 2013.

[18] B. LEPRI ET AL., *Fair, Transparent, and Accountable Algorithmic Decision-making Processes*, Philosophy and Technology, 31, 611–627, 2018.

[19] M. ANANNY, K. CRAWFORD, *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*, New Media & Society, 20(3), 973-989, 2018.

[20] H. THIMBLEBY, *Explaining code for publication*, Journal of Software: Practice and Experience, 33, 975–1001, 2003.

[21] T. GEBRU ET AL., *Datasheets for datasets*, Communications of the ACM, 64(12), 86-92, 2021.

[22] M. BRUNDAGE ET AL., *Toward trustworthy AI development: Mechanisms for supporting verifiable claims*, ArXiv. /abs/2004.07213, 2020.

[23] H. SHAH, *Algorithmic accountability*, Philosophical Transactions of the Royal Society A, 376:20170362.20170362, 2018.

[24] L. FLORIDI, *Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical*, Philosophy and Technology, 32, 185–193, 2019.

[25] A. BUHMANN, J. PASMANN, C. FIESELER, *Managing algorithmic accountability: balancing reputational concerns, engagement strategies, and the potential of rational discourse*, Journal of Business Ethics, 163(2), 265–280, 2020.

[26] L. FLORIDI, *Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions*, Philosophical Transactions of the Royal Society A, 374:2083.20160112, 2016.

[27] D. DANKS, A.J. LONDON, *Algorithmic Bias in Autonomous Systems*, In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI 2017), 2017.

[28] N.T. LEE, *Detecting racial bias in algorithms and machine learning*, Journal of Information, Communication and Ethics in Society, 16(3), 252-260, 2018.

[29] B. GIOVANOLA, S. TIRIBELLI, *Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms*, AI & society, 38, 549–563, 2023.

[30] S. BAROCAS, M. HARDT, A. NARAYANAN, *Fairness in machine learning*, NIPS Tutorial, 2017.

[31] M. VEALE, R. BINNS, *Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data*, Big Data & Society, 4(2), 2053951717743530, 2017.

[32] S. MILLER, *Machine learning, ethics and law*, Australasian Journal of Information Systems, 23, 2019.

[33] B. TARRAN, *What can we learn from the Facebook—Cambridge Analytica scandal?*, Royal Statistical Society, 2018.

[34] D.J. SOLOVE, & P.M. SCHWARTZ, *Information privacy law*, Aspen Publishing, 2020.

[35] W. HOLMES, D. BEKTIK, B. WOOLF & L. ROSE, *Ethics in AIED: Who cares?*, In S. Isotani, E. Millán, A. Ogan, P. Hastings, & R. Luckin (Eds.), Artificial Intelligence in Education, 11625, 424–425, Cham, Switzerland: Springer, 2019.

[36] K. PORAYSKA-POMSTA, G. RAJENDRAN, *Accountability in human and artificial intelligence decision-making as the basis for diversity and educational inclusion*, In J. Knox, Y. Wang, & M. Gallagher (Eds.), Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices, 39–59, Singapore: Springer, 2019.