Interaction Design and Architecture(s) Journal - IxD&A, N.64, 2025, pp. 63 - 79
DOI: 10.55612/s-5002-064-002sp

Trustworthy Learning Analytics for Smart Learning Ecosystems

Barbi Svetec, Blaženka Divjak

University of Zagreb, Faculty of Organization and Informatics, Pavlinska 2, Varaždin, Croatia
{barbi.svetec, blazenka.divjak}@foi.unizg.hr

Abstract. Even though the benefits of learning analytics (LA) have been recognized in research, there are still challenges to its more widespread adoption, which could contribute to the development of smart learning ecosystems beyond 2030. Trustworthiness of LA is an aspect which could play an important role in the adoption of LA, especially in the artificial intelligence (AI) era. Different dimensions of trustworthiness of LA have been researched, but the framework is still fuzzy. We conducted a scoping literature review to provide more clarity, and in this paper, we presented an overview of theoretical considerations and dimensions of trustworthiness of LA. We grouped the identified dimensions into social and technological aspects, and pointed to horizontal dimensions related to both. These dimensions were used as the basis for a comprehensive definition of the trustworthiness of LA. Finally, we identified the challenges and open questions related to using LA to support smart learning ecosystems.

Keywords: learning analytics, trustworthiness, trust, data, ethics, data security, algorithms, fairness

1 Introduction

In order to respond to the needs of the contemporary society, the Timisoara Declaration [1] called for the development of *smart* learning ecosystems. In this context, it emphasized the increase in data production in *phygital* environments, and called for the use of data for monitoring, to foster transparency, reveal problems, expectations and needs in learning processes, providing grounds for “an increase in individuals’ and ecosystems’ *smartness*”. Furthermore, it pointed to the value of drawing questions from data and statistical analyses, drew attention to the use of trace data, and stressed the obligation to provide interoperable, easy-to-use analytics tools. In other words, the Declaration highlighted the importance of meaningful learning analytics (LA) in the development of smart learning ecosystems.

For over a decade, LA has been increasingly used in teaching and learning practice to enable better understanding of learning processes, provide grounds for targeted learning support, enhance learning experiences, and foster self-regulated learning. However, regardless of the recognized opportunities it provides in taking teaching and learning forward, LA is still not widespread and used in its full potential in higher education (HE) [2], let alone other levels of education, where research on LA has been much less substantial [3]. Reasons for that may be various, including the lack of policies and strategic support from academic leadership for the implementation of LA, poor capacities for data-informed decision-making, and differing stakeholder engagement [4].

Moreover, the adoption of LA can be significantly affected by another factor: trust [2]. Therefore, the aspect that has been recently gaining more and more attention in the LA arena is the trustworthiness of LA. Logically, without being able to trust LA, stakeholders (primarily students and teachers, but also institutions and decision-makers) cannot be expected to engage in LA or use it in decision-making related to teaching and learning. LA applications should, therefore, not only be innovative and useful, but also trustworthy, to support making data-informed decisions about learning [5]. Notably, it is not only about LA tools: trust in LA also involves the perspective of trust in stakeholders [6].

A number of studies have been conducted touching upon or more deeply analyzing the issues related to trust in LA, often drawing on the existing ethical frameworks. However, as pointed out in some recent work, *trust* has generally been explored without taking a critical look at its elements and definition [7] [8]. So, what does the notion of trustworthiness of LA actually entail? What are the main challenges? What are the essential components of building a trustworthy LA system? How can we support more trustworthy LA? Finally, what questions do we still need to answer to be able to harness the potentials of trustworthy LA in developing smart learning ecosystems?

Our aim is to refer to the existing research, explore the overall concept and group the dimensions of trustworthiness of LA. Based on that, we will propose a definition of trustworthiness of LA. Furthermore, we will identify the key challenges and open questions. Providing clarity may facilitate the setting up of trustworthy LA systems in the future and, consequently, their more widespread adoption, enhancing the development of smart learning ecosystems.

2 Background

Even though LA and artificial intelligence (AI) are distinctive research fields [9], it has been pointed out that, as a multi-disciplinary field, LA involves AI [10], and that LA is "increasingly unthinkable without AI" [8]. Building on the Timisoara Declaration, which called for meaningful use of data and analyses in the development of smart learning ecosystems, the Troyes Declaration [11], in the light of recent developments, drew special attention to AI. Particularly, in the context of people-centered smart learning ecosystems beyond 2030, it stressed the potentials of a collaborative interplay between humans and AI.

Trustworthiness is a prominent issue in the context of AI, so much that, as an example from the European context, the European Commission has developed guidelines to promote it [12], followed by a set of specific guiding documents, including guidelines for the use of AI in teaching and learning [13]. Therefore, while exploring trustworthy LA, we will start by briefly referring to the components of trustworthy AI.

According to the mentioned guidelines [12], trustworthy AI has three components: it should be lawful, ethical and robust. It requires human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing; and accountability. The said guidelines have often been referred to in the current research on trustworthy or trusted LA. However, there is a complex relationship between AI and LA, and it is worthy to consider whether LA might need a more specific framework.

When considering LA in particular, there has been some research aiming to identify the aspects of trustworthy LA. For example, a criteria catalogue has been proposed [5] aiming to support students in assessing the trustworthiness of LA applications, covering seven core areas: autonomy; protection; respect; non-discrimination; responsibility and accountability; transparency; and privacy and good data governance.

Moreover, to facilitate trusted implementation of LA, a checklist has been proposed [14] with action points to be considered by managers and decision-makers, including: determination (why do you want to apply LA?), explain (be open about your intentions and objectives), legitimate (why are you allowed to have the data?), involve (involve all stakeholders and data subjects), consent (make a contract with the data subjects), anonymize (make the individual not retrievable), technical (procedures to guarantee privacy), external (if you work with external providers).

There have also been initiatives to promote policy development supporting responsible and ethical adoption of LA, like the Jisc's Code of Practice for LA [15], the Open University's Policy on Ethical Use of Student Data for LA [16], and the SHEILA Framework [17].

A recent study [8] established the "contours of trust" in LA and steps that higher education institutions (HEIs) could take to foster it. The study identified, inter alia, the elements of trust in institutional understanding of students' learning (competency, transparency, understanding of learning theory, institutional values, privacy and ethics, data quality, finance/staff resource, student as individual); elements impacting the trustworthiness of data (data completeness/relevance, ethical framework, competence in data analysis, data stewardship, transparency, quantification, understanding context, purpose); elements of trust in LA planning and design (stakeholder involvement, understanding of purpose, appropriate methodology/theory base, cohesive vision/strategy, understanding of student success, data availability/appropriateness, ethical framework, transparency); and elements of trust in operationalization of LA (transparency regarding purpose, data capture and use, problem/context understanding, consent, staff/student capabilities/training, understanding limitations and benefits of LA, key stakeholder involvement, organizational readiness, clear ethical and legal framework).

However, while research abounds, and frameworks and guidelines [18] have been developed for ethical implementation of LA, there is still a lack of their practical implementation [19]. Moreover, these have not been envisaged as generic, conceptual frameworks defining trustworthy LA and its dimensions.

3 Research questions and methods

To provide an overview of the topics researched and discussed in relation to trustworthy LA, a preliminary scoping literature review has been conducted. Scoping reviews are used "to clarify working definitions and conceptual boundaries of a topic or field" [20]. They can serve, inter alia, to identify key characteristics related to a certain concept, and as a step before a systematic literature review [21].

The aim of this scoping literature review has not been to synthesize the evidence and results reported in the literature, but to provide insights into the dimensions of trustworthiness of LA and the related open questions, relevant with respect to the development of smart learning ecosystems. Therefore, the review focuses on the following research questions (RQ):

1. What are the dimensions of trustworthiness of LA?
2. How can we define the trustworthiness of LA, taking into account the identified dimensions?
3. What are the open questions related to the trustworthiness of LA?

The review started by extracting the initial group of research papers from two major bibliographical databases: Web of Science (WoS) and Scopus. The search was done on 25 July 2022, based on the following search string: ("trustworthy" OR "trustworthiness" OR "trust") AND "learning analytics". To provide a comprehensive overview, no time limitation was imposed on the search. The search identified the total of 75 papers in WoS and 73 in Scopus. Among the papers identified in Scopus, 21 were not already included in WoS, so only these were included in further analysis. The search was repeated in WoS on 15 February 2024. This time, only papers published since July 2022 were considered, resulting in 26 additional WoS entries.

In both phases, titles and abstracts were then manually screened. Entries not representing research papers, such as conference proceedings summaries, were excluded from further analysis. All the research papers were then checked against two general criteria: whether they deal with LA and, if yes, whether they deal with trust or trustworthiness. After detailed examination, the final list of 23 articles was extended with additional 14, using the snowballing technique, i.e. using the reference lists and citations in the extracted papers to identify additional papers [22]. Additionally, three other documents (legal, reports) were included in the final review. The final analysis, therefore, included 40 papers (marked with an asterisk in the reference list).

4 Results and discussion

4.1 Dimensions of trustworthy LA (RQ1)

The research papers included in the review dealt, in more or less detail, with fostering trust in LA or ensuring the trustworthiness of LA. Here, it should be noted that, at least for the purpose of this paper, trust and trustworthiness are considered as different concepts. While trust is more subjective, seen as a subjective belief that LA can be trusted, trustworthiness is perceived as a more objective quality of an LA system. In fact, in a recent study on the trustworthiness in LA [8], trust has been defined as "subjective, psycho-social, relational and often asymmetrical and founded on the character/values/credibility and track record/consistency/expertise of the person/organization requiring our trust". The study states that "the level of trust can be influenced by the transparency of the process/requirements and predictability of the envisaged outcomes, and a belief in fairness and benevolence".

The fact that users trust an LA system does not necessarily mean that the system is actually trustworthy, nor does a trustworthy LA system necessarily get the trust of its users. We acknowledge, though, that the two are mutually closely connected, and therefore this review takes both aspects into account when talking about trustworthiness.

The review of the selected research papers shows that the trustworthiness of LA can be considered through two aspects: social (ethical and legal) and technological (data, algorithms and infrastructure), or "soft" and "hard". The two should by no means be considered in silos, as they are mutually interrelated and often dependent. Each consideration encompasses a range of more specific dimensions. There are also horizontal dimensions, important in relation to both the social and technological aspects.

Social aspects. Even though there is an understanding that insights into learners' behaviors provided by LA are valuable to learners, teachers, as well as educational providers, the collection and use of data for LA have been facing various ethical challenges, including those related to privacy, informed consent, and de-identification of data, location and interpretation, as well as classification, management and storage of data [23]. A systematic literature review of ethical considerations covered in empirical research analyzed the available research from the perspective of the following areas of ethics: transparency, privacy, informed consent, responsibility, minimizing adverse impacts, validity and enabling interventions [24].

These ethical considerations are closely related to regulation. In the European Union (EU), data privacy, protection and security have been strictly regulated by the General Data Protection Regulation (GDPR) [25]. Protection principles outlined in GDPR include, inter alia, lawfulness, fairness and transparency, integrity, confidentiality, and accountability. The EU AI Act [26] includes some applications of AI in education among high-risk AI systems, particularly when it comes to the use of AI in evaluating learning outcomes, including those used to steer learning processes, and monitoring and detecting students' prohibited behavior during tests, opening up new questions related to LA. Furthermore, there are institutional rules and regulations regarding the ethical use of data [16].

In the following sections we provide a more detailed overview of the social aspects of trustworthy LA. It should be noted that responsibility and accountability will be discussed separately, as they are seen as more horizontal.

Privacy, data protection and data security. LA systems frequently collect sensitive data, including on demographics, grades, and interactions with online content [27]. How students' data privacy is regulated depends much on institutional data governance, as well as data security, consent and accountability [28]. There are issues related to how students' personal data are collected, stored, analyzed, and presented to various stakeholders. In particular, according to a recent SLR, privacy and data protection-related issues in LA include collecting sensitive and excessive data in the data collection phase; anonymization, storing sensitive data and calculation problems in the data analysis phase; misuse of data, anonymization, storing sensitive data and calculation problems in the data reporting phase. Furthermore, issues cross-cutting the different phases include those related to the definition of LA privacy, communication and transparency, power relationships, lack of quality and stakeholders' conservative attitudes, and issues related to legislation [29]. Although students' personal data are often anonymized when processed and stored, it has been pointed out that this is but a small step in ensuring a comprehensive educational data governance for LA [30], and "data security is a poor stand-in for privacy" [31]. A recent SLR found there are privacy and data-protection related differences in the attitudes of stakeholders, including those between students, between different stakeholders (students and teachers, students and developers), and between students and reality [29].

It has been pointed out that students trust their HEIs in terms of respecting their privacy, not disclosing and misusing their data, but using the data in a way that serves student interests [32]. However, students also often have concerns about privacy, control of what data and for what purposes are collected, implications of the analyses [16], and the storage and accessibility of their data later on [33]. Research has found they are more cautious about sharing certain types of data, especially their personal information and data trails of their online behavior [30]. It has been found that students' privacy concerns, risk and control perceptions, are related to their trust and non-self-disclosure behaviors, and that reduced perception of risk may contribute to the adoption of LA [34]. Similarly, teachers also have concerns about access rights, transparency and consent, as well as data sharing and ownership, especially when it comes to digital tools providers outside education institutions [6].

Agency, autonomy and control. It has been observed that the relationship between the privacy principles and the acceptance and use of LA points to the need to actively involve students, as well as other stakeholders (including teachers, administrators, instructional designers) in the implementation of LA [30]. There have been many calls to engage end-users in LA development [35], including educating students about institutional LA practices [36], giving them more agency in relation to how data is used in LA [14], or including teachers in customizing multimodal LA solutions [35], which can affect their sense of agency and trust in LA outputs. Nevertheless, despite the calls to give students more agency in how data is used in LA, their concerns about privacy may still prevent students from taking part in LA [37].

Students' agency is related to their consent to participate in LA, which demonstrates respect and supports their autonomy and voluntary collaboration [37] [23]. However, it has been recommended that students should be given control which exceeds the possibility of opting out. They should have access to their records, rationales for decisions based on the collected data, and be able to add to the collected data [16], as well as correct or remove information [38] as part of the data governance. They should be informed and included whenever their data, either individual or aggregated, is in question [39]. When it comes to control, the issue of ownership of LA data has been pointed out as essential, specifically the question of who owns the data derived from the raw data collected from users [27], especially relevant once third-party providers come into the scene. Unbalanced power distribution in data processing and decision-making in LA has been pointed out as an obstacle to promoting trust in LA [2].

Furthermore, besides autonomy, concerns about surveillance, also expressed by students themselves [16], have been characterized as contradictory to educational values [2]. For example, data collection has been found to threaten students' agency [40], and there are concerns about personalized approaches leading to spoon-feeding instead of empowering students [41].

To enable trust, LA should be developed in alignment of priorities of different stakeholders, in order to support not only inclusiveness, but also relevance of LA with respect to the needs of end-users. In this context, human-centered LA has been gaining prominence [42], and various co-design models have been proposed including different stakeholders in the design, development and implementation of LA [2]. Research suggests the need to include students in collecting, analyzing and using their data [16]. However, although a number of LA-related studies have called for more inclusive strategies, LA is still often developed without including the key stakeholders (primarily students and teachers) in the process [43]. End-users often just provide data and get the results of analyses [35].

Stakeholders. Trustworthiness of LA is also related to the issue of trust in stakeholders (HEIs, third parties), in terms of competence (experience and expertise) to implement LA or deal with ethics-related issues. This relates to, for example, competence in e-learning systems or data analysis, as well as awareness of the context. Moreover, when it comes to third parties, there are concerns about commercial targeting, and data ownership and sharing. [6] In this sense, the issue of *AI loyalty* has been pointed out, and the question of who AI systems actually work for, related more to the ethics of the people than of the technology. It has been noted that, to increase transparency and trustworthiness, AI systems should be aligned with the interests of students and others affected, which again points to the involvement of stakeholders in the design and implementation of AI [44], applicable also to LA.

Technological aspects. Privacy and ethical use of data are closely linked to trust in algorithms and systems using student data. Research has been dealing with infrastructure protecting student privacy, at the same time maintaining the utility of LA methods, as well as algorithms and methods enabling us to measure and mitigate potential risks, including algorithmic fairness. [19]

Technological considerations are, in various ways, related to the social ones. Importantly, when considering the legal perspective, in the EU, the GDPR [25] states that data protection principles have to be considered in the design of any product or activity. Protection principles outlined in the GDPR include, inter alia, storage limitation, and integrity and confidentiality, that includes the encryption of personal data. The EU AI Act points out the obligation to guarantee the right to privacy and personal data protection through all parts of an AI system's lifecycle [26].

Data, algorithms, accuracy and fairness. The accuracy, reliability and fairness of LA is very much related to the data it is based on. However, there are a number of factors that can affect the data and introduce bias to LA results, consequently also leading to biased algorithms.

In terms of data accuracy, teachers have expressed concerns about outdated data, different data channels, lack of data governance, power relationships, response bias, and technical problems [6]. Furthermore, to have a comprehensive view of teaching and learning processes in blended environments, various data from digital and physical contexts would be needed, calling for the application of multimodal LA. Differing data sources can affect the accuracy, relevance, interpretability and actionability of LA, which can, in turn, influence users' sense of trust and agency. [35] Furthermore, it has been pointed out that insights from LA are “objective to the extent that the reliability of observed patterns can be enhanced by a sample that is large in volume and diverse in scope” [2].

Moreover, students’ consent and opting out can influence the predictive power of LA models. It has been shown that biases exist in predictive models, which is partly caused by samples which are not representative, and restricted data may lead to reduced accuracy of machine-learned models. [37] There has been a concern that students opting out from LA, which can differ among student subpopulations (e.g., demographic groups), may skew predictive models and introduce bias [37]. Besides inconsistencies in student data, psychological factors affecting student data have also been recognized as issues related to the precision of prediction [6].

Another concern related to the reliability of LA is related to “fake learners”. This specifically refers to Massive Open Online Courses (MOOCs), where LA can be biased by users who take advantage of anonymity and openness of MOOCs, e.g., by using multiple accounts or unauthorized collaboration. However, it has been found that the behavior of “fake learners” differs from that of “true learners”, and according to some research, LA results may differ depending on whether “fake learners” are removed from the data or not. This points to the need for research methods that are more verifiable, robust and generalizable [45] [46]. On another point, if they know they are being observed, students and teachers can start acting in a way that “satisfies algorithms” [2].

Algorithmic fairness, although in the focus in the fields of machine learning and AI, has only recently got more prominence in the field of LA [19] [47]. It has been pointed out that LA aims to support all students, but potential bias and lack of impartiality of algorithms may lead to inaccurate modelling for student groups which are not well represented, affecting the fairness of LA [37].

Research has shown a possibility of bias of LA models against students of some demographic subgroups, and there are unknowns related to which unfairness mitigation algorithm or metric to use. In relation to this, it has been found that preserving fairness does not necessarily negatively affect LA utility, in fact, ensuring fairness can sometimes enhance utility. Interestingly, it has also been found that data bias does not necessarily lead to predictive bias [47].

Finally, the complex structures and non-transparent decision-making mechanisms of machine learning models make it hard for users to interpret and, consequently, trust the predictions [48]. It has been noted that trustworthy AI in education relates to the interpretability of prediction outcomes and processes, as well as the inclusion of subject experts in the development of prediction models [48]. In this context, the need for “explainable AI” has been emerging [49].

Infrastructure. For successful LA infrastructure, one of the essential factors is the ability to decide on the necessary software architecture concept: there are different architectural concepts in LA and the selection of the appropriate one is not straightforward. [50]

LA systems, especially when meant to serve a large number of users, face challenges such as large amounts of rapidly arriving data, possibly from diverse sources (various virtual environments, as well as multimodal data), which are often duplicated, fragmented, non-standardly represented and differently identified. [50] Such challenges have to be faced by big data software infrastructure, while proactively anticipating the volume, velocity and variety of the incoming data [51]. Importantly, big data should also be seen from the perspective of its veracity, meaning its trustworthiness and accuracy [50]. However, it should be noted that many institutional LA system use data that are far from big data, and in their case, these challenges are not so prominent.

Privacy frameworks and regulations have important implications for the design of tools, architectures and practices in LA [52]. For example, the abovementioned GDPR principle of protection by design poses multiple challenges for the development of LA architecture [50], and the EU AI Act sets out clear requirements and obligations related to high-risk AI systems [26].

Accessibility. Besides data on students with disabilities often being poorly represented, there is also the issue of interface design and LA that is accessible and usable for diverse students, especially when it comes to visualizations and dashboard design [53] [54]. It is important to ensure that LA tools and systems can be accessed and used by different groups of users, including those with disabilities (e.g., visually impaired students, or those with learning disorders like dyslexia), as well as those speaking minority or foreign languages.

Horizontal aspects. Besides the described dimensions, which can be generally grouped as social or technological, there are also dimensions (such as transparency and accountability, also emphasized in the GDPR [25]) which should be horizontally applied to both.

Transparency. Transparency has been identified as an important driver for successful implementation of LA [30]. Some research indicates that students may be comfortable with personal data collection if it led to more effective support, which may suggest that transparency has an important role in students' perceptions of privacy in LA. Students should know what data are collected, by whom, for what purposes and benefits, and who will have access to the data; they should also have access to the analyses and related feedback. [16] Furthermore, it has been found that students may be willing to share more data if an LA system provides "rich and meaningful information" [30].

The importance of considering transparency in the design of LA and AI-based algorithms has also been pointed out [55]. This implies that, for LA to be widely accepted, its AI components need to be interpretable, so that the reasons for AI systems' outputs can be understood. Moreover, they should be explanatory, meaning that AI systems provide information behind the reasoning [55], with more and more focus being put on *explainable AI* [49].

DOI: 10.55612/s-5002-064-002sp

Responsibility and accountability. To support trust in LA, committed leadership and sound policy is needed [2], at the institutional, as well as higher levels, calling for responsibility and accountability. Some authors have noted that responsibility and accountability overlap with transparency, privacy, good data governance and autonomy [5]. Research suggests that (HE) institutions should be mindful of providing transparency, student control, and clarifying the consequences, benefits, risks and possible biases to students [16], and are "obliged to educate" their students about institutional LA practices [36]. Moreover, HEIs have been described as "information fiduciaries", with the responsibility to use students' data to promote students interests and the institution's educational mission [32]. Recent research has also found that teachers trust the competence of HEIs and usefulness of LA, but not so much third parties when it comes to privacy and ethics, or data accuracy because of issues like a lack of data governance or outdated data [6]. Again, this links to the issue of *AI loyalty* [44] and working in line with institutional values and the best interest of students in mind.

Moreover, institutional leaders and teachers should also be accountable for the implementation of LA, for example, providing more transparency on the use of data [37]. Some research has found that students inherently trust their institutions and expect them to use their data ethically [16]. Furthermore, to support trust in AI, teachers' knowledge about AI should be enhanced, and supported by professional development [56]. In this respect, in broader terms, institutional leaders should be responsible to support the development of digital, AI and data literacy of teachers, so that LA results are understandable to them and do not represent black boxes. More particularly, teachers should be provided with support and learning opportunities to acquire skills needed to use LA and extract relevant information from data [6]. Curricula should also foster the AI and data literacy of students, to enable the understanding of LA and self-regulation of learning.

Finally, it has been found that institutional trust, relation between the amount of data collection and perceived benefits, and comfort regarding teachers' use of data for learning engagement are essential for students' decisions on participation in LA [37].

An overview. As can be seen from the described aspects and dimensions, they are mutually related, often dependent, and should therefore be seen as parts of a whole, as presented in Figure 1.

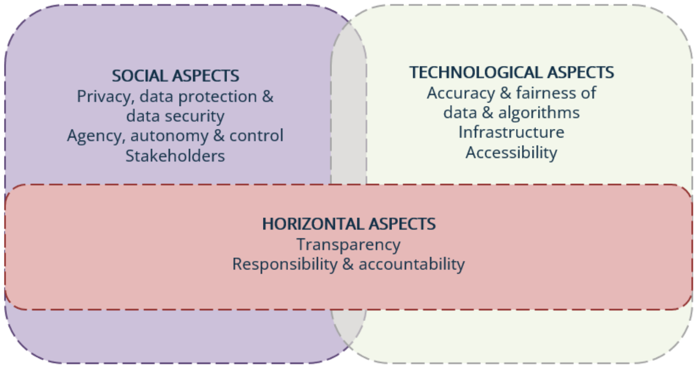

Fig. 1. Aspects and dimensions of trustworthy LA

4.2 Definition of trustworthiness of LA (RQ2)

As it seems that issues related to the trustworthiness of LA have been researched and discussed in various contexts, but the concept itself has not been clearly defined, considering the identified dimensions, we will propose a definition of trustworthiness of LA.

To start with, let us consider the generic definition of the term *trustworthiness*. According to Oxford Learner's Dictionaries, *trustworthiness* means the quality of always being good, honest, sincere, etc. so that people can rely on you. Interestingly, the dictionary identifies *reliability* as a synonym. However, in linguistics, absolute synonymy has often been questioned, and it has been pointed out that "natural languages abhor absolute synonyms just as nature abhors a vacuum" [57]. Therefore, looking beyond linguistics, we should consider the definition of *reliability* in the contexts of research in general, and in particular data science. Considering this, it should be noted that *reliability* is closely related to the described technological considerations (either of data, algorithms, or analysis results), and as such, is only one aspect of *trustworthiness*.

Taking all of this into account, we propose to define the trustworthiness of LA *as the quality of LA which abides by legal rules and ethical principles related to learners' privacy, their data security and control, is based on non-biased data and algorithms, transparently used, and can be trusted to support all learners in successful acquisition of learning outcomes.*

4.3 Open questions (RQ3)

This scoping literature review provided an overview of aspects and dimensions which have been researched and discussed and should be considered when developing trustworthy LA. However, this was not a comprehensive systematic literature review nor empirical research and should be seen as a basis for further research attempts to clarify the concept of trustworthy LA and what it entails.

The review also flagged some of the possible issues related to the mutual connection and interdependence between the identified dimensions, but it did not empirically investigate their mutual relationships. In future research, this could be researched in more detail, through empirical research with relevant stakeholders.

The described dimensions and related issues lead to a number of questions, which we will open in the following paragraphs, as a possible basis for future research and policy discussion.

The issues of autonomy, voluntary collaboration, and consent to take part in LA are closely connected to the issue of data bias, algorithmic bias, and consequently, LA accuracy. As stated in some previous work, student data does not have to be "personally identifiable", however, more granular and identifiable data also lead to more opportunities for analytics [31]. On the one hand, we need to ensure these ethical prerequisites to support students' trust in LA. On the other hand, how can anyone trust LA if it is based on non-comprehensive data? Students and teachers can also learn to act in a way that satisfies algorithms [2]. Furthermore, students are often reluctant about the data on their learning behavior being exploited [30], but this gives us irreplaceable insights. How can balance be found in this respect? What is the perspective for machine-generated, synthetic data [58]?

Additionally, would such balance be enabled by the existing legal frameworks and other relevant regulation? Even though firm regulations are essential from the data protection perspective, how do they reflect on the possibilities for further development in the area of LA? Does firm regulation limit the possibilities? What are the implications of the overarching AI regulation at the EU level? Would more flexibility bring more value to LA users?

The strategic planning done by educational leadership should consider the trustworthiness of LA when planning the systemic implementation of LA. Are educational leaders sufficiently familiar with the requirements of trustworthy LA and the related regulatory framework? Are they aware of teachers' and students' major obstacles to using the benefits of LA? These questions are related to the wider strategic planning at the institutional or system level. If there is no strategic approach to the development and implementation of LA, important issues related to trustworthiness may be left to developers, designers or tool-providers with limited strategic perspectives, including vendors outside of HE systems.

From the institutional perspective, it is also important to ensure that both students and teachers understand LA, in order to be able to use it meaningfully. How to ensure that, on the one hand, LA is explainable, and on the other hand, that students and teachers have the prerequisites to understand it? In this respect, it is important to consider the need to revise curricula and provide professional development opportunities. How to support other stakeholders to understand LA?

From the policy perspective, when it comes to the responsibility for the implementation of LA, what are the essential areas for investment? Should more focus be put on social or technological aspects? At this point, is it more important to focus on teachers' competences, digital and data literacy, or should we focus on infrastructure? What comes first?

Furthermore, it is important to consider that not all stakeholders – including students, teachers, educational leaders, institutions and systems, researchers, designers and developers – have equal perspectives, for example, on privacy and data protection [29].

So, are their requirements in terms of trustworthy LA the same? What are the differences? Who has the priority? If we consider the proposed definition of trustworthiness of LA, and that is based on learner-centered approaches, the priority should be given to learners' needs. Is that so in practice? In this respect, it would be worth further investigating the perspectives of students and teachers within the same institution.

Some groups of users seem to be more concerned about privacy issues than others. Moreover, while some students are motivated by competition and comparisons with peers, others find it demotivating [59]. Among different cultures and systems, there are also differences in terms of more collectivist or individualist orientation – for example, some cultures see HE as more public, whereas other see it a private good. Is it possible to adopt a one-size-fits-all approach to the trustworthiness of LA? Or is it essential to consider differences between educational traditions and cultures, including those specific to an institution or a subject-area? Can the perception of trustworthiness be culturally colored?

Finally, what is the role of sound learning design in ensuring the trustworthiness of LA? How can we ensure that LA results are didactically explainable and can be meaningfully used to support pedagogical improvements?

Answering these questions is important in order to streamline the use data to enhance teaching and learning processes and support the development of smart learning ecosystems.

5 Conclusion

We presented an overview of aspects and dimensions of trustworthiness of learning analytics (LA) based on a literature review. We identified two general aspects of trustworthiness of LA: social (ethical and legal) and technological (data, algorithms and infrastructure), or "soft" and "hard". The two aspects should by no means be considered in silos, as they are mutually interrelated and often dependent. We also recognized there are horizontal dimensions, transparency, and responsibility and accountability, which should be taken into account throughout the process of strategic planning and implementation of trustworthy LA.

Based on that, we proposed a definition of trustworthy LA, explaining the *trustworthiness* of LA as *the quality of LA which abides by legal rules and ethical principles related to learners' privacy, their data security and control, is based on non-biased data and algorithms, transparently used, and can be trusted to support all learners in successful acquisition of learning outcomes.*

We also identified a number of open questions related to the implementation of trustworthy LA, as one of the pillars of smart learning ecosystems. The major questions are related to the role of leadership in the strategic implementation of LA, ensuring the explainability of LA, the dichotomy between the regulatory framework and exploiting the full potential of LA, as well as between privacy concerns and enabling non-discriminatory results. Furthermore, we open questions related to (possibly varying) stakeholder perspectives, as well as possible cultural differences leading to diverse perceptions of trustworthy LA.

Acknowledgement. This work has been supported by the Croatian Science Foundation under the project *Trustworthy Learning Analytics and Artificial Intelligence for Sound Learning Design* – TRUELA (IP-2022-10-2854).

CRediT author statement. **Barbi Svetec:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review and editing, Visualization. **Blaženka Divjak:** Conceptualization, Methodology, Writing – review and editing, Visualization, Supervision, Funding acquisition.

References

1. ASLERD *et al.*, Timisoara declaration: 'BETTER LEARNING FOR A BETTER WORLD Through People Centered Smart Learning Ecosystems,' 2016. Accessed: April 10, 2024. [Online]. Available: https://www.aslerd.org/summary-of-timisoara-declaration/ https://doi.org/10.55612/s-5002-064-001d)
2. Tsai Y. S., Whitelock-Wainwright A., and Gašević D.: More Than Figures on Your Laptop: (Dis)trustful Implementation of Learning Analytics, *Journal of Learning Analytics*, vol. 8, no. 3, pp. 81–100 (2021), doi: 10.18608/jla.2021.7379.
3. de Sousa E. B. G., Alexandre B., Ferreira Mello R., Pontual Falcão T., Vesin B., and Gašević D.: Applications of Learning Analytics in High Schools: A Systematic Literature Review, *Frontiers in Artificial Intelligence*, vol. 4 (2021) doi: 10.3389/frai.2021.737891.
4. Tsai Y.-S. and Gasevic D.: Learning analytics in higher education --- challenges and policies, in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, New York, NY, USA: ACM, pp. 233–242 (2017) doi: 10.1145/3027385.3027400.
5. Veljanova H., Barreiros C., Gosch N., Staudegger E., Ebner M., Lindstaedt S.: Towards Trustworthy Learning Analytics Applications: An Interdisciplinary Approach Using the Example of Learning Diaries, pp. 138–145 (2022) doi: 10.1007/978-3-031-06391-6_19.
6. Alzahrani A. S., Tsai Y. S., Aljohani N., Whitelock-wainwright E., Gasevic D.: *Do teaching staff trust stakeholders and tools in learning analytics? A mixed methods study*, vol. 71, no. 4. Springer US (2023) doi: 10.1007/s11423-023-10229-w.
7. de Fine Licht K. and Brülde B.: On Defining 'Reliance' and 'Trust': Purposes, Conditions of Adequacy, and New Definitions, *Philosophia*, vol. 49, no. 5, pp. 1981–2001 (2021), doi: 10.1007/s11406-021-00339-1.
8. Slade S., Prinsloo P., and Khalil M.: 'Trust us,' they said. Mapping the contours of trustworthiness in learning analytics, *Information and Learning Science*, vol. 124, no. 9–10, pp. 306–325 (2023), doi: 10.1108/ILS-04-2023-0042.
9. Rienties B., Køhler Simonsen H., and Herodotou C.: Defining the Boundaries Between Artificial Intelligence in Education, Computer-Supported Collaborative Learning, Educational Data Mining, and Learning Analytics: A Need for Coherence, *Frontiers in Education*, vol. 5, (2020), doi: 10.3389/feduc.2020.00128.
10. Chatti M. A., Dyckhoff A. L., Schroeder U., Thüs H.: A reference model for learning analytics, *International Journal of Technology Enhanced Learning*, vol. 4, no. 5/6, p. 318 (2012), doi: 10.1504/IJTEL.2012.051815.
11. ANP *et al.*: Troyes Declaration (Timisoara 2.0): Better Learning for a Better World 2.0: People Centred Smart Learning Ecosystems beyond 2030 (2024). Accessed: Oct. 1, 2024. [Online]. Available: https://www.aslerd.org/troyes-declaration/ https://doi.org/10.55612/s-5002-064-002d
12. E. Commission: Ethical Guidelines for Trustworthy AI (2019).

13. E. Commission: Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Brussels (2022) doi: 10.2766/153756.
14. Drachsler H. and Greller W.: Privacy and analytics - it's a DELICATE issue a checklist for trusted learning analytics, *ACM International Conference Proceeding Series*, vol. 25-29-Apri, pp. 89–98 (2016), doi: 10.1145/2883851.2883893.
15. Sclater N.: Developing a Code of Practice for Learning Analytics, *Journal of Learning Analytics*, vol. 3, no. 1 (2016), doi: 10.18608/jla.2016.31.3.
16. Slade S., Prinsloo P., and Khalil M.: Learning analytics at the intersections of student trust, disclosure and benefit, *ACM International Conference Proceeding Series*, pp. 235–244 (2019), doi: 10.1145/3303772.3303796.
17. Tsai Y.-S. S. *et al.*: The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics, *Journal of Learning Analytics*, vol. 5, no. 3, pp. 31–47 (2018), doi: 10.18608/jla.2018.53.2.
18. Slibar B., Gusic Mundjar J., Rako S., and Simic D.: Co-occurrence patterns of issues and guidelines related to ethics and privacy of learning analytics in higher education—literature review, in *LAK22: 12th International Learning Analytics and Knowledge Conference*, New York, NY, USA: ACM, pp. 577–582 (2022), doi: 10.1145/3506860.3506974.
19. Ladjal D., Joksimović S., Rakotoarivelo T., and Zhan C.: Technological frameworks on ethical and trustworthy learning analytics, *British Journal of Educational Technology*, vol. 53, no. 4, pp. 733–736 (2022), doi: 10.1111/bjet.13236.
20. Peters M. D. J., Godfrey C. M., Khalil H., McInerney P., Parker D., Soares C. B.: Guidance for conducting systematic scoping reviews, *International Journal of Evidence-Based Healthcare*, vol. 13, no. 3, pp. 141–146 (2015), doi: 10.1097/XEB.0000000000000050.
21. Munn Z., Peters M. D. J., Stern C., Tufanaru C., McArthur A., Aromataris E.: Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach, *BMC Medical Research Methodology*, vol. 18, no. 1, p. 143 (2018), doi: 10.1186/s12874-018-0611-x.
22. Wohlin C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering, in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering - EASE '14*, New York, New York, USA: ACM Press (2014) pp. 1–10. doi: 10.1145/2601248.2601268.
23. Slade S. and Prinsloo P.: Learning Analytics: Ethical Issues and Dilemmas, *American Behavioral Scientist*, vol. 57, no. 10, pp. 1510–1529 (2013), doi: 10.1177/0002764213479366.
24. Cerratto Pargman T. and McGrath C.: Mapping the Ethics of Learning Analytics in Higher Education: A Systematic Literature Review of Empirical Research, *Journal of Learning Analytics*, vol. 8, no. 2, pp. 123–139 (2021), doi: 10.18608/jla.2021.1.
25. E. Union: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da) (2016) *Official Journal of the European Union*.
26. E. Union: Artificial Intelligence Act (2024), doi: https://artificialintelligenceact.eu/the-act/.
27. Pardo A. and Siemens G.: Ethical and privacy principles for learning analytics, *British Journal of Educational Technology*, vol. 45, no. 3, pp. 438–450 (2014), doi: 10.1111/bjet.12152.
28. Prinsloo P., Slade S., and Khalil M.:The answer is (not only) technological: Considering student data privacy in learning analytics, *British Journal of Educational Technology*, vol. 53, no. 4, pp. 876–893 (2022), doi: 10.1111/bjet.13216.
29. Liu Q. and Khalil M.: Understanding privacy and data protection issues in learning analytics using a systematic review, *British Journal of Educational Technology*, vol. 54, no. 6, pp. 1715–1747 (2023), doi: 10.1111/bjet.13388.

30. Ifenthaler D. and Schumacher C.: Student perceptions of privacy principles for learning analytics, *Educational Technology Research and Development*, vol. 64, no. 5, pp. 923–938 (2016) doi: 10.1007/s11423-016-9477-y.
31. Jones K. M. L. *et al.*: Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios, *portal: Libraries and the Academy*, vol. 23, no. 3, pp. 485–515 (2023), doi: 10.1353/pla.2023.a901565.
32. Jones K. M. L., Rubel A., and LeClere E.: A matter of trust: Higher education institutions as information fiduciaries in an age of educational data mining and learning analytics, *Journal of the Association for Information Science and Technology*, vol. 71, no. 10, pp. 1227–1241 (2020), doi: 10.1002/asi.24327.
33. Adejo O. and Connolly T.: Learning Analytics in a Shared-Network Educational Environment: Ethical Issues and Countermeasures, *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 4 (2017) doi: 10.14569/IJACSA.2017.080404.
34. Mutimukwe C., Viberg O., Oberg L., and Cerratto-Pargman T.: Students’ privacy concerns in learning analytics: Model development, *British Journal of Educational Technology*, vol. 53, no. 4, pp. 932–951 (2022), doi: 10.1111/bjet.13234.
35. Rodríguez-Triana M. J., Prieto L. P., Martínez-Monés A., Asensio-Pérez J. I., Dimitriadis Y.: The teacher in the loop: Customizing multimodal learning analytics for blended learning, *ACM International Conference Proceeding Series*, pp. 417–426 (2018), doi: 10.1145/3170358.3170364.
36. Jones K. M. L. *et al.*: ‘We’re being tracked at all times’: Student perspectives of their privacy in relation to learning analytics in higher education, *Journal of the Association for Information Science and Technology*, vol. 71, no. 9, pp. 1044–1059 (2020), doi: 10.1002/asi.24358.
37. Li W., Sun K., Schaub F., Brooks C.: Disparities in Students’ Propensity to Consent to Learning Analytics, *International Journal of Artificial Intelligence in Education*, vol. 32, no. 3, pp. 564–608 (2022), doi: 10.1007/s40593-021-00254-2.
38. Prinsloo P. and Slade S.: Ethics and Learning Analytics: Charting the (Un)Charted, in *Handbook of Learning Analytics*, Society for Learning Analytics Research (SoLAR) pp. 49–57 (2017), doi: 10.18608/hla17.004.
39. Prinsloo P. and Slade S., Student privacy self-management, in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, New York, NY, USA: ACM, pp. 83–92 (2015), doi: 10.1145/2723576.2723585.
40. Rubel A. and Jones K. M. L.: Student privacy in learning analytics: An information ethics perspective, *The Information Society*, vol. 32, no. 2, pp. 143–159 (2016), doi: 10.1080/01972243.2016.1130502.
41. Tsai Y.-S., Perrotta C., and Gašević D.: Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics, *Assessment & Evaluation in Higher Education*, vol. 45, no. 4, pp. 554–567 (2020), doi: 10.1080/02602938.2019.1676396.
42. Buckingham Shum S., Ferguson R., and Martinez-Maldonado R.: Human-Centred Learning Analytics, *Journal of Learning Analytics*, vol. 6, no. 2 (2019), doi: 10.18608/jla.2019.62.1.
43. Selwyn N. and Gašević D., The datafication of higher education: discussing the promises and problems, *Teaching in Higher Education*, vol. 25, no. 4, pp. 527–540 (2020), doi: 10.1080/13562517.2019.1689388.
44. Council of Europe: Artificial Intelligence and Education (2022).
45. Alexandron G., Ruipérez-Valiente J. A., Lee S., Pritchard D. E.: Evaluating the Robustness of Learning Analytics Results Against Fake Learners, pp. 74–87 (2018), doi: 10.1007/978-3-319-98572-5_6.
46. Alexandron G., Yoo L. Y., Ruipérez-Valiente J. A., Lee S., Pritchard D. E.: Are MOOC Learning Analytics Results Trustworthy? With Fake Learners, They Might Not Be!,

International Journal of Artificial Intelligence in Education, vol. 29, no. 4, pp. 484–506, (2019), doi: 10.1007/s40593-019-00183-1.

47. Deho O. B., Zhan C., Li J., Liu J., Liu L., Duy Le T.: How do the existing fairness metrics and unfairness mitigation algorithms contribute to ethical learning analytics?, *British Journal of Educational Technology*, vol. 53, no. 4, pp. 822–843 (2022), doi: 10.1111/bjet.13217.
48. Duan X., Pei B., Ambrose G. A., Hershkovitz A., Cheng Y., Wang C.: Towards transparent and trustworthy prediction of student learning achievement by including instructors as co-designers: a case study, *Education and Information Technologies* (2023)=, doi: 10.1007/s10639-023-11954-8.
49. Brdnik S., Podgorelec V., and Šumak B., Assessing Perceived Trust and Satisfaction with Multiple Explanation Techniques in XAI-Enhanced Learning Analytics, *Electronics (Switzerland)*, vol. 12, no. 12 (2023), doi: 10.3390/electronics12122594.
50. Ciordas-Hertel G. P., Schneider J., Ternier S., Drachsler H.: Adopting trust in learning analytics infrastructure: A structured literature review *Journal of Universal Computer Science*, vol. 25, no. 13, pp. 1668–1686 (2019).
51. Taylor S. and Munguia P.: Towards a data archiving solution for learning analytics, in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, New York, NY, USA: ACM, pp. 260–264 (2018), doi: 10.1145/3170358.3170415.
52. Hoel T., Griffiths D., and Chen W.: The influence of data protection and privacy frameworks on the design of learning analytics systems, in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, New York, NY, USA: ACM, pp. 243–252 (2017), doi: 10.1145/3027385.3027414.
53. Chen W.: Learning analytics for inclusive higher education, in *Proceedings 28th International Conference on Computers in Education , 2020*, Asia-Pacific Society for Computers in Education, pp. 113–118 (2020).
54. Khalil M., Slade S., and Prinsloo P.: Learning analytics in support of inclusiveness and disabled students: a systematic review," *Journal of Computing in Higher Education* (2023), doi: 10.1007/s12528-023-09363-4.
55. Kloos C. D. *et al.*, H2O Learn - Hybrid and Human-Oriented Learning: Trustworthy and Human-Centered Learning Analytics (TaHCLA) for Hybrid Education, in *2022 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, pp. 94–101 (2022), doi: 10.1109/EDUCON52537.2022.9766770.
56. Nazaretsky T., Ariely M., Cukurova M., and Alexandron G.: Teachers' trust in AI-powered educational technology and a professional development program to improve it," *British Journal of Educational Technology*, vol. 53, no. 4, pp. 914–931 (2022), doi: 10.1111/bjet.13232.
57. Cruse D. A.: *Lexical Semantics*. Cambridge University Press (1986).
58. Savage N.: Synthetic data could be better than real data, *Nature* (2023), doi: 10.1038/d41586-023-01445-8.
59. Divjak B., Svetec B., and Horvat D.: Learning analytics dashboards: What do students actually ask for?, in *LAK23: 13th International Learning Analytics and Knowledge Conference*, New York, NY, USA: ACM, pp. 44–56 (2023), doi: 10.1145/3576050.3576141.