Where Does All the Data Go? A Review of Research on E-Assessment Data

Michael Striewe[a]
University of Duisburg-Essen, Essen, Germany

Keywords: Technology-enhanced Assessment, Data, Literature Review.

Abstract: E-Assessment systems produce and store a large amount of data that can in theory be interesting and beneficial for students, educators and researchers. While there are already several reviews that elicit commonly used methods as well as benefits and challenges, there is less research about the various contexts and forms in which data from e-assessment system is actually used in research and practice. This paper presents a structured review that provides more insights into the contexts and ways data and data handling is actually included in current research. Results indicate an emphasis on some contexts in current research and that there are two dimensions of data usage.

1 INTRODUCTION

Giving feedback to students and judging their progress in terms of marks or grades is an integral part of learning processes in almost every kind of educational setting. In recent decades, it became increasingly common to support diagnostic, formative and summative assessments with technology-enhanced assessment systems; shortly called e-assessment systems. Reasons for using such systems include but are not limited to automation of grading, creation of opportunities for self-regulated learning, and the need to conduct assessments in the time of social distance. In any of these cases, e-assessment systems will collect and store a large amount of mostly personal data, mainly about students and the interaction with the system. In addition, they will also produce data such as grades.

While there is usually a clear reason why to use an e-assessment system at all, it is sometimes less clear why data is stored or where and when it is used. There are some obvious cases, e. g. when data is used to create grades and feedback. Research on educational technology also needs empirical data that can be taken from e-assessment systems. Data can also be used to discover exam fraud, to predict course outcomes, or to improve the quality of an exercise, a course or a curriculum. However, the theoretical option to use data for one of these purposes is not enough to collect large amounts of personal data. In some cases, aggregated or anonymous data may be sufficient, while consent to use data is required in other contexts.

There is a large body of research in the area of educational data mining and learning analytics that has been published in recent years. Several literature reviews have been published in that area, too, both for the general case (e. g. (Aldowah et al., 2019; Peña-Ayala, 2014; Romero and Ventura, 2010) and for specific domains of study (e. g. (Ihantola et al., 2015)). More recently than learning analytics, also the field of assessment analytics has been defined (Ellis, 2013), but detailed reviews of research do not yet exist for that area.

The focus of available reviews is primarily on methods and approaches (and sometimes also tools) or on the particular information gains that can be produced by these methods. These reviews thus provide valuable answers to the question *how* and *why* to use data from e-assessment systems *in a particular context*. However, they do not explicitly answer the question *where* and *when* such data is actually used at all and thus do not tell anything about *all existing contexts*. The latter can partly be concluded from the contexts for which the former type of reviews exist, but that will obviously miss any research for which no such review has been created yet. This is even more likely, as the use of "big data" in education has also been seen critical due to the risk of social exclusion and digital dividedness in recent years (Timmis et al., 2016) and thus approaches to data usage other than

[a] https://orcid.org/0000-0001-8866-6971

Striewe, M.
Where Does All the Data Go? A Review of Research on E-Assessment Data.
DOI: 10.5220/0010879700003182
In *Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022) - Volume 1*, pages 157-164
ISBN: 978-989-758-562-3; ISSN: 2184-5026

educational data mining and learning analytics may have emerged.

This paper asks two research questions to get a clearer picture on the use of data and measures for data handling in current research on e-assessment systems: (RQ 1) In what contexts is data from e-assessment systems used? (RQ 2) In what way and when is data from e-assessment systems used? Notably, the goal of this paper is to identify contexts and dimensions of data usage that have not yet been covered by detailed reviews – it is not the goal to synthesise learnings from all of these context. A detailed exploration of any of the discovered contexts is subject to future research.

The remainder of this paper is organized as follows: Section 2 presents the methodology used for the review of current research. Section 3 presents results and thus answers to both research questions. Section 4 discusses further consequences from the results as well as threads to validity. Section 5 concludes the paper.

2 METHODOLOGY

The first step in this research was to perform a simple, systematic search in the SCOPUS[1] literature database. The search used the terms *(e-assessment OR "computer aided assessment" OR "technology enhanced assessment") AND data*. Additional terms like *automatic assessment* have been tested as well, but produced too much irrelevant results from areas other than educational assessment.

Search terms were applied to titles, abstracts and keywords of the papers in the database. In addition, the search was limited to articles, conference papers, book chapters, and reviews. A first run of the search was performed in March 2021, followed by a second run in October 2021 to include more up-to-date results. Results are only reported for the total set of search results throughout the remainder of this paper. Since data for 2021 cannot be complete before the end of the year, only papers published in 2020 or earlier are considered.

2.1 Exclusion Criteria and Classification

The second step in this research was to classify all papers based on their content. Classifications were recorded by assigning at most two labels to each paper, denoting the primary topics of each paper. Papers that did not cover the topic of data usage in or from technology-enhanced assessment or even were not concerned with educational assessment at all were labeled with label *"off"*. These papers are excluded from any further investigations. Five additional labels were used for papers that covered technology-enhanced assessment, but did not focus on data usage:

- Label *"study on e-assessment"* was used for papers that discussed studies on e-assessment where data was not taken from the e-assessment systems, but solely from other sources such as surveys or interviews. A different label *"data use in study"* was used for papers that discussed studies on e-assessment where data was indeed taken from the e-assessment systems and that were thus considered relevant for the review.
- Label *"system design"* was used for papers that only presented and discussed system design but not specifically data handling. A different label *"data handling"* was used for papers in which system design was discussed with an emphasis on data handling, which was considered relevant for the review.
- Label *"review"* was used for papers that presented reviews of other publications but did not originally report on the use of data within some system.
- Label *"theory on e-assessment"* was used for papers that discussed general and abstract theories on e-assessment or process models for e-assessment, but did not discuss the actual use of data.
- Label *"domain-specific item handling"* was used for papers that are concerned with domain-specific e-assessment in domains that involve the term "data", such as "data structures" in the context of computer science.

For the remaining papers, at most two labels were assigned the describe best the primary kind of data or means of data handling contained in that paper. There was no pre-defined list of possible labels before starting the review, but new labels were defined as necessary. To decide which label(s) can be applied to a paper, title and abstract were read first. In most cases, these contained sufficient information to select one or two appropriate labels. In case of doubt, the full paper was read to make a decision.

Figure 1 provides an overview on the classification process.

2.2 Data Analysis

Labels have then been analysed to answer the first research question. In particular, papers with the same

[1] https://www.scopus.com/

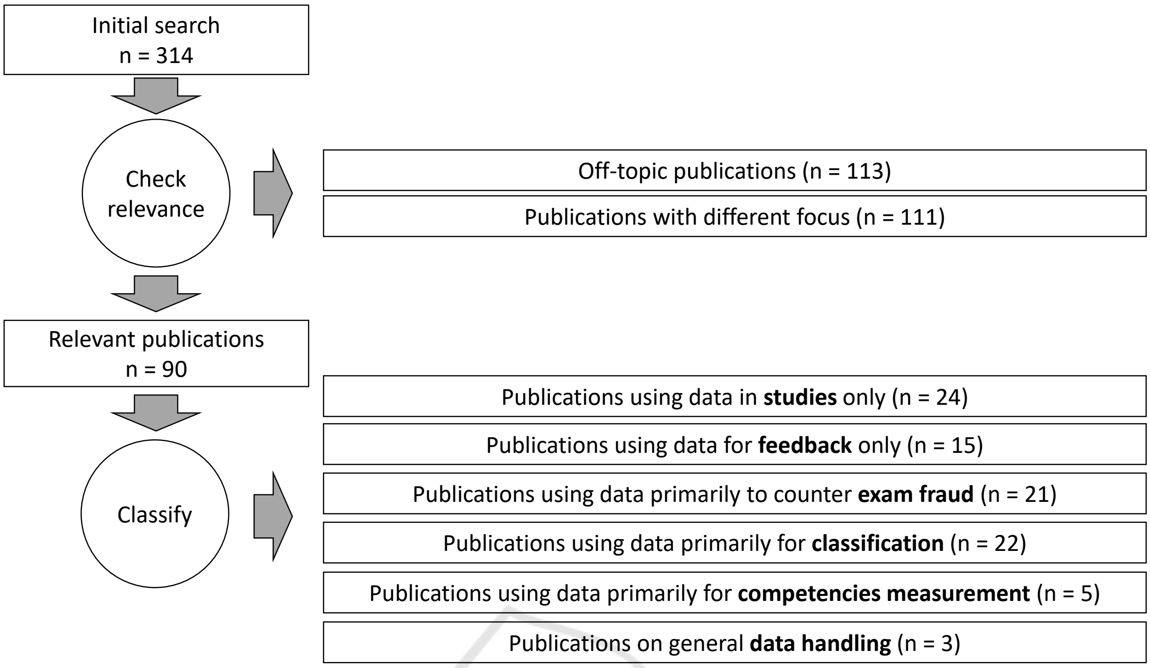

Figure 1: Overview on the classification process applied in this literature review.

or similar labels have been clustered into groups that represent specific aspects of data usage. Papers from these groups have then been analysed in more detail to answer the second research question.

3 RESULTS

The search returned a total amount of 314 publications. From those, 224 publications were excluded with the labels listed in the left-hand part of table 1. The remaining 90 papers were assigned with labels as listed in the right-hand part of the same table.

3.1 Bibliometric Data

Publication years (see figure 2) seem to reveal an increasing interest in the topics that are considered relevant for the review. The oldest publication is from 2005 and thus fairly new (at least in comparison to the oldest excluded search result which is from 1948) and there were not more than two relevant publications per year until 2010. Different to that, there were at least 6 relevant publications per year since 2014 and more than 50% of all relevant publications have been published within the last four years.

The most frequent publication venue among the relevant papers is the IEEE Global Engineering Education Conference (EDUCON) with five papers. It is followed by two journals (British Journal of Educational Technology and International Journal of Emerging Technologies in Learning) and two proceedings series (Lecture Notes in Computer Science and Lecture Notes on Data Engineering and Communications Technologies), each with four publications.

3.2 Contexts and Forms of Data Usage

The 90 relevant publications that were included in the study can be divided into several groups. These will be discussed in the following paragraphs in decreasing order of their size.

The largest group contains 24 publications that report on some research study on e-assessment, where at least a part of the data used in the study comes directly from an e-assessment system. In comparison to the large body of research in the field of technology-enhanced assessment this seems to be a small number. It may thus not be representative, but only a random sample that contains all search terms by chance. Nevertheless, it can be concluded that using data from an e-assessment system for research purposes is at least a very relevant use case, if not indeed the most frequent one. Common to most of the studies is that data is usually extracted once. The focus of research is usually not an individual person, so that anonymous or aggregated data can be used. However, persons must remain identifiable if data from e-assessment systems should be combined with data from other sources such as interviews or questionnaires.

Table 1: Number of papers per label for excluded papers (left) and included papers (right). Note that papers from the right-hand part could be labels with more than one label and thus the sum of all labels is larger than the total amount of papers.

Label	Papers	Label	Papers
off	113	data use in study	24
study on e-assessment	46	feedback	18
system design	44	plagiarism	8
theory on e-assessment	9	authentication	8
review	6	quality	8
domain-specific item handling	6	privacy	7
		improvement	6
		prediction	6
		dishonesty	5
		competency measurement	5
		adaptivity	5
		classification	5
		data handling	3

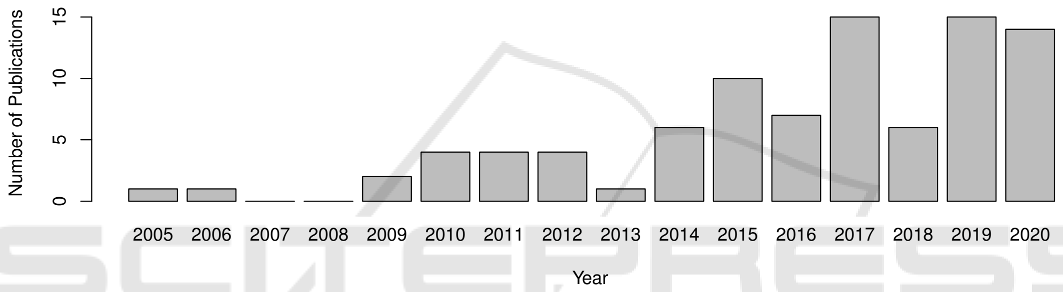

Figure 2: Number of papers per year that have been included in the analysis.

A major group of 21 papers is concerned with measures to detect or prevent exam fraud. A significant share of the publications in that category origin from the recent "TeSLA"-project[2]. Both aspects are related closely to each other, but can be distinguished by the way they use data: One aspect is the *detection* of plagiarism and other forms of academic dishonesty. A total amount of 11 publications tackle that topic and discuss approaches on how to detect dishonesty from e-assessment data. The topic is not specific to technology-enhanced assessment but also relevant for paper-based exams. However, technology-enhanced assessments make it easier to analyse solutions (cf. e. g. (Opgen-Rhein et al., 2019)) and to collect additional information like keystroke characteristics (Baró et al., 2020) to reveal dishonesty. At the same time, unproctored e-assessments may make it easier for students to commit exam fraud (cf. e. g. (Amigud et al., 2018)). With the respect to data usage, anonymous data can be sufficient to check for indicators for exam fraud and personal data will only be revealed in conjunction with actually suspicious cases. Some mechanisms combine data from e-assessment systems with other data (e. g. previous submissions of coursework or resources from the internet (Bañeres et al., 2019)). Moreover, checks can be run once (e. g. after an exam) and need no constant access to data.

The other aspect is authentication and privacy, which can be used to *prevent* exam fraud. It is discussed by a total amount of 14 publications. The important trade-off here is how much personal data must be revealed to ensure a proper authentication and how much data can be kept anonymous (Okada et al., 2019; Muravyeva et al., 2019). Different to the previous aspect, data is usually used continuously (e. g. to make sure that the person who logged in for an exam or enrolled for a course is indeed the person that works on the exam/course) and in conjunction with external data sources (e. g. for single sign-on mechanisms).

An almost equally large group of 22 publications is concerned with data classification approaches that employ mathematical or statistical models. There is a wide range of application areas: Adaptive e-assessments (e. g. (Runzrat et al., 2019; Birjali et al., 2018; Geetha et al., 2013)), prediction of exam results or completion time (e. g. (Carneiro et al., 2019;

[2] https://cordis.europa.eu/project/id/688520/

Usman et al., 2017; Gamulin et al., 2015)), or quality measurement and improvement (e. g. (Stack et al., 2020; Azevedo et al., 2019; Derr et al., 2015)). Pure classification can also be used as a means of data aggregation for feedback generation (Nandakumar et al., 2014; Sainsbury and Benton, 2011). Besides adaptive e-assessment, these aspects are mentioned in the papers several times in conjunction with the term "learning analytics". Different to the research studies on e-assessment mentioned above, the focus of the publications is not on a detailed measurement that is performed once in the context of academic research, but on continuous or frequent use of data for the respective purposes.

Although the mathematical or statistical methods might be similar for different purposes, the kind of data is not. Adaptivity clearly requires continuous use of individual data, since such systems adapt their contents based on individual responses while learners are working on an assessment or assignment. In contrast to that, prediction is usually performed frequently and can also involve data from other sources such as learning management systems. Data used for quality measurement and improvement is usually aggregated and anonymous, while data that should help students to improve their way of learning in personalized systems clearly needs to be related to that person (Saul and Wuttke, 2014).

The next group contains 15 publications that are concerned solely with giving feedback on individual items (while there are two papers that are not only concerned with feedback, but also with classification and another paper that tackles feedback and plagiarism). Since 15 papers is a rather low number given the fact that most e-assessment systems are primarily designed to give feedback, these papers can hardly be considered representative for the way in which e-assessment data is used for feedback generation. Nevertheless, these papers already show that feedback generation requires continuous usage of data. If feedback is solely directed towards the learners, feedback mechanisms can use anonymous data. Feedback for teachers that reports about a larger group of learners can use aggregated data, but feedback in exams obviously is related directly to individual, identifiable persons.

A special aspect of feedback generation is competency measurement, for which 5 publications could be discovered that are explicitly related to that topic. Similar to the low number of papers on general feedback generation, it is possible that much research on that topic is published without direct relation to e-assessment and has thus not been discovered by the search terms used for this simple survey. An interesting aspect with respect to data handling is the fact, that competency measurement not only uses data from e-assessment systems, but also produces data (i. e. measured competency levels) that may be stored as additional data directly associated with individual persons in some kind of learner model (Bull et al., 2012; Florián et al., 2010).

Finally, 3 publications discuss general topics of data handling within e-assessment systems independent of a particular use case. The discussions cover meta-data management (Sarre and Foulonneau, 2010), conversion between data formats (Malik and Ahmad, 2017) and approaches to data visualization (Miller et al., 2012).

3.3 Dimensions of Data Usage

Besides a classification into topics, the survey also helps to identify characteristics of data usage along different dimensions. One dimension that was already mentioned above is the frequency of data use. Data can be extracted from an e-assessment system once for single use, i. e. in the context of a research study. It can also be extracted or used frequently. This is the case for example when solutions are checked for plagiarism at the end of an exam or when data is extracted at the end of a course for quality assurance. Finally, data can also be used continuously, e. g. for adaptivity, competency measurement, or during authentication.

Another dimension is the granularity and richness of data. For many studies or for quality assurance it is sufficient to use anonymous or aggregated data that does not reveal too much individual details. Also grading and feedback generation can often be performed without revealing personal data of the answer's author. Anonymous data is particularly beneficial with respect to data privacy. Aggregated data is more compact to handle than detailed data and thus e. g. easier to visualize. Other scenarios like competency measurement or adaptivity nevertheless require individual, identifiable data since they concern individual students. Using anonymous or aggregated data is not possible in that case, although that means to involve more sophisticated algorithms to handle large amounts of data and to ensure data privacy. In very specific cases, particularly in conjunction with extensive research studies, but also for prediction, plagiarism checks or some ways of authentication it may be necessary to combine e-assessment data with data from other sources. That can be achieved by contributing data to a general data repository. The resulting data is very rich and detailed, but also very sensitive with respect to data privacy.

An overview on the two dimensions of data usage and some scenarios is given in table 2.

4 DISCUSSION

The results show that there are various views on e-assessment data that all get remarkable attention in current research. While it is nearby that data from e-assessment systems is used in research study and can be used in the context of learning analytics or adaptivity, it is interesting to see that there is also an emphasis of research for the complex topic of dishonesty and privacy. This adds a new legal and ethical perspective to the established perspectives of educational and technical aspects in research on e-assessment systems. As expected in the reasoning for conducting this study, it also reveals research about data from e-assessment systems that is not related to "classical" perspectives of educational data mining or learning analytics.

It is probably due to the way the literature search was performed that the classical educational perspective of feedback generation and competency measurement seems to be underrepresented in the search results. At the same time, a purely technical perspective that solely focuses on data handling appears even less often. This allows for the interpretation that research has an emphasis on the purpose of data usage rather than the way of data handling.

Notably, no time constraint was used during the literature search and most papers have been published fairly recently. Given the fact that e-assessment systems are known for much longer, writing about data usage or handling appears to be a relatively new topic that currently gets increased interest. One reason for that could be an increasing awareness for privacy issues that makes it necessary to justify why and which data is collected and stored. At the same time, mathematical and statistical methods seem to become more usable and thus make it more appealing to perform data analyses in large scale.

The latter aspect is also supported by the fact that the results of the survey are similar to the results of a broader survey on artificial intelligence applications in higher education (Zawacki-Richter et al., 2019): In that survey, profiling and prediction was identified as a major use case (58 out of 146 studies), followed by assessment and evaluation (36 studies), intelligent tutoring (29 studies), and adaptivity and personalisation (27 studies). Hence only the aspect of authentication, privacy and dishonesty was not covered in that survey, which is not surprising as these topics are usually not associated with the use of artificial intelligence. That again stresses the point that it is not sufficient to take the perspective of methods, but also the perspective of contexts.

4.1 Threads to Validity

The search in the SCOPUS database returned 314 papers from which more than 100 were completely off-topic and a similar share did at least not match the core of the review criteria. The resulting number of papers appears to be quite low in comparison to the large amount of research on e-assessment that has been published in recent years. This is probably due to the fact that the key term "data" is not always included in the title or abstract of these papers and they are thus not included in the search results. Thus, there is a risk that some topic might have gained attention in research, but has never been published in a way that it was included in the search results. At the same time, there is no reason to assume that this problem actually applies to a significant amount of research topics, since the review already covers a wide range of aspects.

Labels have been assigned to the papers mainly based on the titles and abstracts. There is a probability that papers might have covered more topics than indicated in these places. Consequently, the review does not include these topics. However, it was not the aim of the review to present a detailed content analysis of all papers, but to identify the main contexts in which data from e-assessment systems is used in what way. It can be assumed that the main purpose of a study is indeed named in the abstract of an paper and thus only minor aspects of some research have not been recorded.

5 CONCLUSION

The review of research presented in this paper achieved to answer both research questions that were stated in the beginning: First, the review identified groups of related topics that have produced remarkable amounts of research and that covered very different perspectives on e-assessment data. Hence, it is now more clear that there are several distinct contexts in which data from e-assessment systems is used for very different purposes. Second, the review shed light on the various forms of data that appears in the different context. From these impressions, two dimensions of data usage could be derived that can be used to classify data usage.

The results from the review can be used in several ways: First, a more detailed content analysis can be

Table 2: Overview on examples for typical scenarios using e-assessment data along the two dimensions of data usage.

Type of data	Cases with one-time use	Cases with frequent use	Cases with continuous use
Anonymous or aggregated data	research studies	quality assurance	feedback
Individual, identifiable data	—	plagiarism check	feedback, adaptivity, competency measurement
Data merged with external sources	research studies	plagiarism check, prediction	authentication

performed for the papers included in the search results to get even more insights into the dimensions of data usage as well as possible interconnections within and between the different contexts of data usage. Second, the results can be used as a starting point to make connections to the usage of other data than e-assessment data in similar context. For example, authentication, privacy and plagiarism may also be relevant topics in other areas of educational technology and beyond, even if academic dishonesty may indeed only be a major problem in the context of assessments. Third, the results can be used to identify research gaps that require further attention. The results so far are surely not yet detailed enough for that purpose, but the fact that e. g. papers on data handling appear relatively rarely in the search results may hint towards the fact that the technical aspects on how to handle data within e-assessment systems could possibly need further attention in future research.

REFERENCES

Aldowah, H., Al-Samarraie, H., and Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and informatics*, (37):13–49.

Amigud, A., Arnedo-Moreno, J., Daradoumis, T., and Guerrero-Roldan, A.-E. (2018). Open proctor: An academic integrity tool for the open learning environment. In *Lecture Notes on Data Engineering and Communications Technologies*, volume 8, pages 262–273.

Azevedo, J., Oliveira, E., and Beites, P. (2019). Using learning analytics to evaluate the quality of multiple-choice questions: A perspective with classical test theory and item response theory. *International Journal of Information and Learning Technology*, 36(4):322–341.

Bañeres, D., Noguera, I., Elena Rodríguez, M., and Guerrero-Roldán, A. (2019). Using an intelligent tutoring system with plagiarism detection to enhance e-assessment. In *Lecture Notes on Data Engineering and Communications Technologies*, volume 23, pages 363–372.

Baró, X., Muñoz Bernaus, R., Baneres, D., and Guerrero-Roldán, A. (2020). Biometric tools for learner identity in e-assessment. In *Lecture Notes on Data Engineering and Communications Technologies*, volume 34, pages 41–65.

Birjali, M., Beni-Hssane, A., and Erritali, M. (2018). A novel adaptive e-learning model based on big data by using competence-based knowledge and social learner activities. *Applied Soft Computing Journal*, 69:14–32.

Bull, S., Wasson, B., Kickmeier-Rust, M., Johnson, M., Moe, E., Hansen, C., Meissl-Egghart, G., and Hammermueller, K. (2012). Assessing english as a second language: From classroom data to a competence-based open learner model. In *Proceedings of the 20th International Conference on Computers in Education, ICCE 2012*, pages 618–622.

Carneiro, D., Novais, P., Durães, D., Pego, J., and Sousa, N. (2019). Predicting completion time in high-stakes exams. *Future Generation Computer Systems*, 92:549–559.

Derr, K., Hübl, R., and Ahmed, M. (2015). Using test data for successive refinement of an online pre-course in mathematics. In *Proceedings of the European Conference on e-Learning, ECEL*, pages 173–180.

Ellis, C. (2013). Broadening the scope and increasing the usefulness of learning analytics: The case for assessment analytics. *British Journal of Educational Technology*, 44(4):662–664.

Florián, B. E., Baldiris, S. M., and Fabregat, R. (2010). A new competency-based e-assessment data model: Implementing the aeea proposal. In *IEEE Education Engineering Conference, EDUCON 2010*, pages 473–480.

Gamulin, J., Gamulin, O., and Kermek, D. (2015). The application of formative e-assessment data in final exam results modeling using neural networks. In *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, pages 726–730.

Geetha, V., Chandrakala, D., Nadarajan, R., and Dass, C. (2013). A bayesian classification approach for handling uncertainty in adaptive e-assessment. *International Review on Computers and Software*, 8(4):1045–1052.

Ihantola, P., Vihavainen, A., Ahadi, A., Butler, M., Börstler, J., Edwards, S. H., Isohanni, E., Korhonen, A., Petersen, A., Rivers, K., Rubio, M. Á., Sheard, J., Skupas, B., Spacco, J., Szabo, C., and Toll, D. (2015). Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies. In *Proceedings of the 2015 ITiCSE Working Group Reports, ITICSE-WGR 2015, Vilnius, Lithuania, July 4-8, 2015*, pages 41–63.

Malik, K. and Ahmad, T. (2017). E-assessment data compatibility resolution methodology with bidirectional data transformation. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(7):3969–3991.

Miller, C., Lecheler, L., Hosack, B., Doering, A., and Hooper, S. (2012). Orchestrating data, design, and narrative: Information visualization for sense- and decision-making in online learning. *International Journal of Cyber Behavior, Psychology and Learning*, 2(2):1–15.

Muravyeva, E., Janssen, J., Dirkx, K., and Specht, M. (2019). Students' attitudes towards personal data sharing in the context of e-assessment: informed consent or privacy paradox? In *Communications in Computer and Information Science*, volume 1014, pages 16–26.

Nandakumar, G., Geetha, V., Surendiran, B., and Thangasamy, S. (2014). A rough set based classification model for grading in adaptive e-assessment. *International Review on Computers and Software*, 9(7):1169–1177.

Okada, A., Whitelock, D., Holmes, W., and Edwards, C. (2019). e-authentication for online assessment: A mixed-method study. *British Journal of Educational Technology*, 50(2):861–875.

Opgen-Rhein, J., Küppers, B., and Schroeder, U. (2019). Requirements for author verification in electronic computer science exams. In *11th International Conference on Computer Supported Education, CSEDU 2019*, volume 2, pages 432–439.

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, 41(4, Part 1):1432–1462.

Romero, C. and Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6):601–618.

Runzrat, S., Harfield, A., and Charoensiriwath, S. (2019). Applying item response theory in adaptive tutoring systems for thai language learners. In *Proc. 11th International Conference on Knowledge and Smart Technology, KST 2019*, pages 67–71.

Sainsbury, M. and Benton, T. (2011). Designing a formative e-assessment: Latent class analysis of early reading skills. *British Journal of Educational Technology*, 42(3):500–514.

Sarre, S. and Foulonneau, M. (2010). Reusability in e-assessment: Towards a multifaceted approach for managing metadata of e-assessment resources. In *5th International Conference on Internet and Web Applications and Services, ICIW 2010*, pages 420–425.

Saul, C. and Wuttke, H.-D. (2014). Turning learners into effective better learners: The use of the askme! system for learning analytics. In *UMAP 2014 Posters, Demonstrations and Late-breaking Results*, volume 1181 of *CEUR Workshop Proceedings*, pages 57–60.

Stack, A., Boitshwarelo, B., Reedy, A., Billany, T., Reedy, H., Sharma, R., and Vemuri, J. (2020). Investigating online tests practices of university staff using data from a learning management system: The case of a business school. *Australasian Journal of Educational Technology*, 36(4).

Timmis, S., Broadfoot, P., Sutherland, R., and Oldfield, A. (2016). Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, 42(3):454–476.

Usman, M., Iqbal, M., Iqbal, Z., Chaudhry, M., Farhan, M., and Ashraf, M. (2017). E-assessment and computer-aided prediction methodology for student admission test score. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(8):5499–5517.

Zawacki-Richter, O., Marín, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(39).