A Systematic Review of the Potential of Machine Learning and Data Science in Primary and Secondary Education

P. Donaldson, N. Ntarmos & K. Portelli

University of Glasgow

October 2017

Table of Contents

Introduction 2
- Data Science and Machine Learning in Education 2
- Suitable Problems in Education 3

Enhancing and Transforming Education through Machine Learning 4
- Curriculum Delivery – Range of Timescales Depending on Complexity 4
- Assessment of Learners- Range of Timescales Depending on Complexity 5
- Curricular Design- Long Term Timescale 5
- Educational Research – Short Term Timescale 6

Overcoming Barriers to Adoption 6
- Lack of Understanding Educational Data Science and its Benefits 6
- Developers Failing to Understand and Meet Educator and Learners Needs 7
- Ensuring Adequate Support for Educational Data Science 7
- Privacy and Ethical Considerations 7

Conclusion 8

References 9

Glossary of Key Terms and Acronyms 13

Technical Appendix 15
- Appendix A- Systematic Review Process 15
- Appendix B- Highly Cited Researchers in Educational Data Science Fields 19
- Appendix C- Research Projects and Organisation Websites for Educational Data Science 23
- Appendix D- Existing Tools, Applications and Products 27
- Appendix E- Complete List of Educational Data Science Documents Analysed 30

Introduction

This review, commissioned by the Royal Society, will examine the current state of the art developments in data science and machine learning within education, if they can be applied more broadly now, or in the future, and the impact they could have on transforming primary and secondary education in the UK. A sample of the total volume of research reviewed is directly referenced in the body of the report to help illustrate key issues. The full list of sources reviewed is included as the last section of the technical appendix. This appendix also provides further details on the systematic review process, important researchers in each sub-field, a list of current and recent research projects and existing tools, applications and products.

Data Science and Machine Learning in Education

The sheer scale and quantity of data generated by modern digital devices poses both significant challenges and new opportunities to better understand and support both learning and learners. Although the term data science was coined in 1974 (Naur, 1974), it's only relatively recently that data science, combined with machine learning approaches, have been specifically applied in support of learning and to address wider issues within the field of education itself.

Data Science is an interdisciplinary area that brings together different people, processes, systems and scientific methods to extract insight from data in various forms. This data can be highly structured with related values such as text, dates or numbers grouped together or unstructured with many individual values that are recorded separately.

Educational data can exist at many different levels of detail and is often spread across several different software systems with learning related events happening over a period of time (Silva and Fonseca, 2017, p. 88). These characteristics combined with the complex, and often messy, nature of learning itself has led to two closely related research communities arising in the form of Educational Data Mining and Learning Analytics (Almosallam and Ouertani, 2014).

Educational Data Mining (EDM) involves "developing, researching and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyse due to the enormous volume of data within which they exist" (Romero and Ventura, 2013, p. 12). While Learning Analytics (LA) "refers to the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance, and spot potential issues" (Johnson et al., 2014).

Although similar sounding, with EDM establishing a distinct identity in 2005 and LA holding its first conference in 2011, they differ in focus with EDM concentrating more on automated methods of both data analysis and adaptation and LA concerned with enhancing teacher judgement and intervention (Baker and Inventado, 2014, p. 62). Both, however, deal with large volumes of data and increasingly rely on complex statistical methods and machine learning techniques.

Machine learning is the name given to a range of different algorithms that can be used by a computer to learn from existing data rather than being specifically programmed to understand it. For example, personal assistants such as Siri, Cortana or Alexa use machine learning algorithms (McTear et al., 2016, p. 1) to first turn our spoken request into text and then to analyse this text to understand what action to carry out. Recent breakthroughs in deep learning such as Alpha Go (Silver et al., 2016) have captured the publics' imagination and brought

seemingly impossible problems for computers to solve, such as driving a car (Teichman and Thrun, 2013), within reach.

Suitable Problems in Education

Formative assessment, or assessment for learning, is recognised as a crucial aspect in improving the learning process but one that requires both learners and teachers to have an accurate understanding of learners' stage of development and appropriate next steps (Wiliam, 2010). However, certain forms of knowledge about the quality of learners thinking and what they need to do next can be difficult for a teacher to develop and act upon with just their own classroom experience (Little, 2007, p. 9).

Both EDM and LA, as part of a broader field of Educational Data Sciences (Smith, 2006), have the potential to help address several issues related to this aspect of learning.

For learners, this could involve:-

- supporting learners own assessment of their current level of understanding. With many automated forms of feedback possible including comparison to an expert model or other learner, feedback highlighting areas of strength and development needs and recommendations about what to study next (Steenbergen-Hu and Cooper, 2013).
- analysing learners' performance in real time or afterwards and recommending follow up activities that offer the right level of challenge to support further progress and development of understanding (Daley et al., 2016).
- deciding when and how to provide feedback when they appear to be having difficulties (Fyfe, 2016).

For educators, useful information would be:-

- fine grained and frequent information about a learner, or classes, rate of progress and what they are likely to be able to learn in the time available (Sergis and Sampson, 2017).
- overviews of particular competencies that the learner has developed and where the current gaps and difficulties lie (Amara et al., 2016).
- recommendations for pairing up learners into groups for collaborative learning based on learners' characteristics and how they operate as a team (Amara et al., 2016).
- details of how difficult particular learning activities were for learners and suggestions for more appropriate activities (Slavuj et al., 2017).

For researchers and instructional designers, they would benefit from:-

- analysing and identifying teacher performance data to better understand the characteristics of effective teaching and learning (Agudo-Peregrina et al., 2014).
- help to uncover more effective sequences of learning and discover hidden relationships and dependencies between certain concepts and skills (Howlin and Lynch, 2014).
- identifying and exploring differences between learners and how big an impact particular forms of learning may have on closing those gaps (Almosallam and Ouertani, 2014).
- discovering new misconceptions in areas of learning, how frequently they appear and possibly whether certain prior learning experiences help to address or accidently reinforce them instead (Ma et al., 2014).
- the connection between different emotional states and the quality of learning and how different forms of feedback affect learners' motivation (Bousbia and Belamri, 2014).

Enhancing and Transforming Education through Machine Learning

Machine Learning, as a family of techniques, is based on probabilistic reasoning rather than absolute certainty of the correctness of a solution which is common in many other computational approaches. This means that difficult problems for computers, such as understanding visual information and the meaning of text and other types of information, are starting to be addressed by researchers but the information generated by these techniques can be less reliable. Currently, they have mostly been applied to trying to understand some aspect of the learning process using data from simple interactions with the computer like keyboard keystrokes or actions with a mouse pointer that are easier to process and interpret. However, it is becoming easier to interpret more complex inputs such as user emotions, gestures, multi-touch inputs and speech which could support the analysis of more complex forms of learning (Blikstein and Worsley, 2016).

One of the big tensions is between machine learning methods that use an explicit theory of learning and those that act more like a black box, such as deep neural networks, where much of the detail about how decisions are made is hidden (Armstrong, 2015). Current findings about Technology to Enhance Learning in general (OECD, 2016, p. 84) point towards the need to use technology to complement, rather than displace, other learning activities mediated by a classroom teacher. For this aim to be realised it's important that both the learner model and the data generated from it is able to be understood by learners and their teachers. Work on how to do this with neural networks created for image recognition is still at a relatively early stage (Zeiler and Fergus, 2013). It could be some time before researchers are able to make the reasoning process behind decisions regarding predictions, or estimations, of learning explicit using similar approaches. However, most Educational Data Science methods currently in use can, and do, incorporate such theories of learning making it much easier for both researchers and users to reason about the decisions being made and the reliability of the information presented (Pena-Ayala, 2014, p. 1436).

Curriculum Delivery – Range of Timescales Depending on Complexity

Adaptive learning platforms have seen some of the biggest investments and quickest transfers of machine learning and data science techniques into widely used educational products. Knewton is currently the biggest and well known of these (Wilson and Zack, 2015) and has been used to build a wide range of interactive learning content from other educational publishers for primary school all the way to university level courses. In the EU, platforms such as Realizeit and Cogbooks also enable other organisations to build adaptive learning content that can make recommendations of what to study next based on how learners perform in assessments and interact with the content in the system.

More ambitious systems currently in the research and development phase include the EU funded Newton project (Alkhaldi et al., 2016) which aims to extend machine learning supported personalisation and automatic recommendations to immersive virtual lab experiences. These will use a richer range of inputs such as multi-touch screens and VR gloves that secondary, college and university learners can access from anywhere.

Assessment of Learners- Range of Timescales Depending on Complexity

There is growing evidence that Computer Based Assessment for Learning (CBAfL) with short elaborative feedback can increase the rate and depth of pupils learning (Shute and Rahimi, 2017). Machine learning techniques can be used to assess the current level of understanding and choose questions with just the right level of difficulty that give the most information about a learner's current stage of development and rate of progress.

Most newer systems that involve some form of learning analytics or educational data mining involve the use of a learning records store (Blanco et al., 2013) that can integrate and store data from a number of different applications. Once this data is gathered, the chosen method of data analysis needs to be applied to generate useful insights about learners and their learning. For more complex methods, like analysing a learner's writing, a model can take many seconds or minutes per essay to analyse and then generate feedback. Even with modern computer servers with multiple processors, the resource requirements rapidly become overly demanding for school based server equipment.

For more complex learner modelling to be feasible, platforms for carrying out the analysis of rich learner data will need to be able to scale the amount of storage and processing power they can use to cope with demand. This will probably involve the use of cloud based services but care will be needed to meet storage requirements which need learners' data to be stored securely and processed in their country of origin. Fortunately, there are some examples of how to construct such platforms using a range of freely available and open-source technology on cloud platforms such as Amazon Web Services (Feild et al., 2016).

Curricular Design- Long Term Timescale

Educational Data Science offers many different opportunities to study curricular designs in terms of the balance of different types of activity in a sequence of learning, how learners move through them and the impact that different types of activity have on engagement or learning gain (Sergis and Sampson, 2017).

Research in this area, which is sometimes referred to as Teaching and Learning Analytics (TLA) is still at a fairly early stage of development with the majority of current studies (52%) focused on classifying designs in terms of the types of elements they contain (Rienties et al., 2015) rather than analysing the flow of learners through a sequence of learning or measuring the relative impact on learning of particular designs (Toetenel and Rienties, 2016, p. 157).

A more mature area is the use of data mining techniques for automatic analysis of learning resource content to help discover, catalogue and then recommend suitable open educational resources that teachers can then use to create sequences of lessons. Commercial systems that take this approach include the American company gooru (www.gooru.org). They also provide a range of classroom level analytics for teachers that show how these resources are being used by learners and provide the ability to remix existing lesson sequences. Within these systems, there is the opportunity to use collaborative filtering to automatically recommend learning resources based on what similar educators have found useful (Klašnja-Milićević et al., 2015). However, this doesn't appear to be a feature that is in wide scale use at the moment.

Educational Research – Short Term Timescale

One of the big challenges when conducting large-scale experiments in schools is how to involve teachers and students in a way that causes minimal disruption to the normal process of learning and teaching but supports randomised controlled experiments. The ASSiSTments system tackles this challenge head on and provides an elegant method for creating an online service with equal benefits for both teachers, learners and educational researchers (Ostrow and Heffernan, 2016). Pre-tests, Post-tests and rich interactive exercises with automatic marking and feedback can be created in a range of different presentation formats which are then randomly assigned to classes participating in research studies. The system provides anonymised learner data that researchers can analyse to identify significant relationships and compare treatments, the teacher receives detailed whole class feedback on how students have performed and the students receive immediate feedback on the work they are completing. Once the experiment is over the most effective versions of the exercises are kept and made available for learners and teachers to use.

A recent independent study by the Standford Research Institute (SRI) showed learners in the 1st year of Secondary School using ASSiSTments for mathematics homework made twice the learning gain at the end of the year compared to schools in the control group on a standardised test (Roschelle et al., 2016). Since 2004 the platform has supported the completion of over 20 randomised controlled studies from a range of different universities with at least another 15 studies in progress. Another pilot study was able to identify that immediate feedback benefitted students with less prior knowledge but made no difference to those with higher prior knowledge helping to close the gap between the two groups (Fyfe, 2016).

Overcoming Barriers to Adoption

Although there are many potential barriers to adoption the four most pressing issues that need to be addressed to realise the full potential of Educational Data Science in transforming learning are:-

1. Lack of understanding about educational data science and its benefits
2. Developers failing to understand and meet educator and learners needs
3. Ensuring adequate support for educational data science
4. Privacy and ethical concerns over the use of learners' data

Lack of Understanding Educational Data Science and its Benefits

A review of learning analytics and educational data mining studies involving learners by Li et al highlighted that the majority of participants (82.9%) were from Higher Education with the rest based in secondary school contexts (17.1%) and none in primary schools (Li et al., 2015). The older field of Intelligent Tutoring Systems (Ma et al., 2014), which has increasingly become integrated within the wider area of learning analytics, fares better in this respect with many trials conducted in both primary and secondary school settings.

Part of the reason for this discrepancy is likely due to the difficulty in reaching the required scale of learning data that is needed for many of the statistical methods that compare learners. This is often referred to as the cold-start problem but research projects such as LEA's box have demonstrated that domain based learning models such as Competence-based Knowledge Space Theory and Formal Concept Analysis can provide an alternative approach that does not require large scale data about learners and work in less technology infused environments (Kickmeier-Rust and Albert, 2016).

Another related issue recently highlighted was the current lack of research studies that try to measure the impact on learners and learning using LA (Ferguson and Clow, 2017). The LACE Evidence hub (Scheffel et al., 2014) at the time the study was carried out contained only 38 pieces of research conducted within a school context and only 15 of them examined whether they helped to improve learning. Support from the wider educational research system, including organisations such as the Educational Endowment Foundation (Educational Endownment Foundation, 2017) will be needed in order to expand both the range and scale of studies.

Developers Failing to Understand and Meet Educator and Learners Needs

An emerging theme from several final project reports, including the NEXT-TELL evaluation (Mayer, 2015) and recently published LEA's toolbox evaluation (Kickmeier-Rust, 2017), indicate that exploring research challenges identified within the learning analytics community is still the main focus. This is understandable given how new Educational Data Science is, however significant figures in the field such as George Siemens (Siemens, 2012) have called for increased collaboration between learning system designers, researchers and educators to help tackle this gap.

Unfortunately, this often means that adoption and wide-scale use of the technology developed doesn't happen in schools as the issues that are being addressed are either a lower priority for teachers and learners or the proof of concept technology itself is awkward to use.

Ensuring Adequate Support for Educational Data Science

Suitable standards could also help to address the issue of the small number of open educational datasets (Koedinger et al., 2010) describing learning interactions that can be used to trial and validate new techniques and approaches before incorporating them into systems for learners. Most of the existing datasets have been used by the Educational Data Mining community but often don't contain a rich enough set of learner interaction data for Learning Analytics purposes. There is also currently a lack of researchers with the set of skills needed to undertake more complex work in Educational Data Science (Buckingham Shum et al., 2013), although the UK does currently have a fairly large proportion of top researchers in this field.

A major initiative currently underway is the Open Learning Analytics platform which was proposed in 2011 by several of the most significant figures in the fields of Learning Analytics and Educational Data Mining (Siemens et al., 2011). This recently received a boost when a prominent developer of learning systems, the Apereo Foundation, announced its support and dedicated resources to help create such a system (Apereo Foundation, 2016). Although aimed at Higher Education institutions the software systems under construction are modular and designed to be integrated into a range of other learning systems through standard interfaces for exchanging data such as the eXperience and IMS Caliper/Sensor Application Programming Interfaces (APIs). Developing a set of standards to describe learner interactions above the level of simple actions such as scores on interactive exercises and time taken will greatly accelerate research and wider scale implementation in this area.

Privacy and Ethical Considerations

A pressing concern with the wide scale adoption of educational data science relates to the collection, use and ownership of learners' data. A recent case involving educational association Stichting Snappet in the EU illustrates some of the difficulties (Griffiths et al., 2016). Pre-installed apps for learning mathematics and languages for primary school children collected

data to classify and predict individual students' success and suggest possible learning interventions to schools. Unfortunately, given the personal and highly sensitive nature of the data, the level of information provided to schools about how the data was to be processed and used was too narrow. Informed consent from parents should also have been obtained prior to data collection but wasn't. Once privacy concerns had been fully addressed they were able to continue but this has not always been the case with the American non-profit inBloom shutdown in 2013 for similar concerns (Kharif, 2014).

If these concerns are to be adequately addressed then privacy and ethical concerns should be built into the design of any system (Métayer, 2010) that uses educational data science to generate insights. Although designed with learning analytics in mind, the DELICATE checklist (Drachsler and Greller, 2016) provides a useful starting point for decision makers considering the use of educational data science. Wherever possible, data should be stored securely, access to it controlled and the data subject should able to see what is stored about them and correct any inaccuracies (Pardo and Siemens, 2014). Aggregation of data into summary values and anonymisation should be used wherever possible when sharing data more widely to reduce the risk of individuals personal information being used in inappropriate ways.

Conclusion

Systematic use of data science and machine learning techniques within regular primary and secondary education contexts is mostly confined to a number of specific technologies and areas of application. Intelligent tutoring systems (Ma et al., 2014), adaptive learning resources (Steenbergen-Hu and Cooper, 2013) and assessments (game based or otherwise) currently have the clearest evidence of helping to improve the rate and depth of learning in primary and secondary school contexts when used as a supplement to enhance classroom instruction. Given the success of ASSiSTments (Ostrow and Heffernan, 2016), the creation of a free interactive exercise platform designed to bring together educational researchers, educators and their students, could provide the largest immediate impact on learning in the UK. It may also encourage the development of data science skills amongst a wider range of educational researchers.

In the near future, work carried out developing predictive models of student risk factors (Hu et al., 2017) could also prove useful in a secondary school context alerting staff earlier to potential issues so that they can prioritise resources where, and when, they could make the biggest impact. Automatic content analysis of learning materials (Daems et al., 2014) and recommender systems (Klašnja-Milićević et al., 2015) could make it much easier for teachers to find, evaluate and use high-quality instructional materials with suitable teacher analytics able to highlight areas of difficulty or where further improvements can be made.

Looking further ahead, advances in cognitive computing mean that processing and understanding complex and richer forms of input such as audio, image and video data (Blikstein and Worsley, 2016) will be feasible. This has the potential to radically transform both the range and type of learner work that could be automatically formatively assessed and analysed to identify larger patterns in learning; giving teachers a deeper insight into the learning process and helping to accelerate the development of teaching expertise and sophisticated use of evidence.

References

Some sources referenced in the report itself are not directly related to Educational Data Science but have instead been used to provide wider context for relevant education or data science and machine learning issues.

Agudo-Peregrina, Á.F., Iglesias-Pradas, S., Conde-González, M.Á., Hernández-García, Á., 2014. Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. Comput. Hum. Behav. 31, 542–550. doi:10.1016/j.chb.2013.05.031

Alkhaldi, T., Pranata, I., Athauda, R.I., 2016. A Review of Contemporary Virtual and Remote Laboratory Implementations: Observations and Findings. J. Comput. Educ. 3, 329–351. doi:10.1007/s40692-016-0068-z

Almosallam, E.A., Ouertani, H.C., 2014. Learning Analytics: definitions, applications and related fields, in: Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013). Springer, Singapore, pp. 721–730. doi:10.1007/978-981-4585-18-7_81

Amara, S., Macedo, J., Bendella, F., Santos, A., 2016. Group Formation in Mobile Computer Supported Collaborative Learning Contexts: A Systematic Literature Review. J. Educ. Technol. Soc. 19, 258.

Apereo Foundation, 2016. Apereo Open Analytics An open platform for optimising learning.

Armstrong, H., 2015. Machines that learn in the wild: Machine learning capabilities, limitations and implications. Lond. Nesta.

Baker, R.S., Inventado, P.S., 2014. Educational Data Mining and Learning Analytics, in: Larusson, J.A., White, B. (Eds.), Learning Analytics. Springer New York, pp. 61–75. doi:10.1007/978-1-4614-3305-7_4

Blanco, Á. del, Serrano, Á., Freire, M., Martínez-Ortiz, I., Fernández-Manjón, B., 2013. E-Learning standards and learning analytics. Can data collection be improved by using standard data models?, in: 2013 IEEE Global Engineering Education Conference (EDUCON). Presented at the 2013 IEEE Global Engineering Education Conference (EDUCON), pp. 1255–1261. doi:10.1109/EduCon.2013.6530268

Blikstein, P., Worsley, M., 2016. Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. J. Learn. Anal. 3, 220–238. doi:10.18608/jla.2016.32.11

Bousbia, N., Belamri, I., 2014. Which Contribution Does EDM Provide to Computer-Based Learning Environments?, in: Peña-Ayala, A. (Ed.), Educational Data Mining, Studies in Computational Intelligence. Springer International Publishing, pp. 3–28. doi:10.1007/978-3-319-02738-8_1

Buckingham Shum, S., Hawksey, M., Baker, R.S., Jeffery, N., Behrens, J.T., Pea, R., 2013. Educational data scientists: a scarce breed, in: Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, pp. 278–281.

Daems, O., Erkens, M., Malzahn, N., Hoppe, H.U., 2014. Using content analysis and domain ontologies to check learners' understanding of science concepts. J. Comput. Educ. 1, 113–131. doi:10.1007/s40692-014-0013-y

Daley, S.G., Hillaire, G., Sutherland, L.M., 2016. Beyond performance data: Improving student help seeking by collecting and displaying influential data in an online middle-school science curriculum. Br. J. Educ. Technol. 47, 121. doi:10.1111/bjet.12221

Drachsler, H., Greller, W., 2016. Privacy and Analytics: It's a DELICATE Issue a Checklist for Trusted Learning Analytics, in: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK '16. ACM, New York, NY, USA, pp. 89–98. doi:10.1145/2883851.2883893

Educational Endownment Foundation, 2017. Educational Endownment Foundation 2015-16 Report [WWW Document]. URL https://educationendowmentfoundation.org.uk/public/files/Annual_Reports/EEF_Annual_Report_2015-16.pdf (accessed 7.10.17).

Feild, J.L., Lewkow, N., Zimmerman, N., Riedesel, M., Essa, A., 2016. A Scalable Learning Analytics Platform for Automated Writing Feedback., in: EDM. pp. 688–693.

Ferguson, R., Clow, D., 2017. Where is the evidence?: a call to action for learning analytics, in: Proceedings of the Seventh International Learning Analytics & Knowledge Conference LAK '17. ACM Press, pp. 56–65. doi:10.1145/3027385.3027396

Fyfe, E.R., 2016. Providing feedback on computer-based algebra homework in middle-school classrooms. Comput. Hum. Behav. 63, 568–574.

Gough, D., Oliver, S., Thomas, J., 2013. Learning from Research: Systematic Reviews for Informing Policy Decisions - A Quick Guide. Alliance Useful Evid.

Griffiths, D., Drachsler, H., Kickmeier-Rust, M., Steiner, C., Hoel, T., Greller, W., 2016. Is Privacy a Show-stopper for Learning Analytics? A Review of Current Issues and Solutions. Learn. Anal. Community Exch. 6.

Howlin, C., Lynch, D., 2014. A framework for the delivery of personalized adaptive content, in: 2014 International Conference on Web and Open Access to Learning (ICWOAL). Presented at the 2014 International Conference on Web and Open Access to Learning (ICWOAL), pp. 1–5. doi:10.1109/ICWOAL.2014.7009203

Hu, X., Cheong, C.W.L., Ding, W., Woo, M., 2017. A systematic review of studies on predicting student learning outcomes using learning analytics. ACM Press, pp. 528–529. doi:10.1145/3027385.3029438

Johnson, L., Adams Becker, S., Estrada, V., Freeman, A., Kampylis, P., Vuorikari, R., Punie, Y., 2014. The NMC horizon report Europe: 2014 schools edition., NMC horizon report. Publications Office, Luxembourg.

Kharif, O., 2014. Privacy Fears Over Student Data Tracking Lead to InBloom's Shutdown.

Kickmeier-Rust, M., 2017. Tools for Competence-Centred, Multi-Source Learning Analytics Final Report (Final Report).

Kickmeier-Rust, M.D., Albert, D., 2016. Theory-driven Learning Analytics and Open Learner Modelling: The Teacher's Toolbox of Tomorrow?, in: UMAP (Extended Proceedings).

Klašnja-Milićević, A., Ivanović, M., Nanopoulos, A., 2015. Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions. Artif. Intell. Rev. 44, 571–604. doi:10.1007/s10462-015-9440-z

Koedinger, K.R., Baker, R.Sj., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J., 2010. A data repository for the EDM community: The PSLC DataShop, in: Handbook of Educational Data Mining.

Li, K.C., Lam, H.K., Lam, S.S., 2015. A Review of Learning Analytics in Educational Research, in: Technology in Education. Technology-Mediated Proactive Learning. Presented at the International Conference on Technology in Education, Springer, Berlin, Heidelberg, pp. 173–184. doi:10.1007/978-3-662-48978-9_17

Little, J.W., 2007. chapter 9 Teachers' Accounts of Classroom Experience as a Resource for Professional Learning and Instructional Decision Making. Yearb. Natl. Soc. Study Educ. 106, 217–240. doi:10.1111/j.1744-7984.2007.00103.x

Ma, W., Adesope, O.O., Nesbit, J.C., Liu, Q., 2014. Intelligent tutoring systems and learning outcomes: A meta-analysis. J. Educ. Psychol. 106, 901–918. doi:10.1037/a0037123

Mayer, H., 2015. NEXT-TELL Project Final Report September 2010 - October 2014.

McTear, M., Callejas, Z., Griol, D., 2016. The Dawn of the Conversational Interface, in: The Conversational Interface. Springer International Publishing, Cham, pp. 11–24. doi:10.1007/978-3-319-32967-3_2

Métayer, D.L., 2010. Privacy by Design: A Matter of Choice, in: Data Protection in a Profiled World. Springer, Dordrecht, pp. 323–334. doi:10.1007/978-90-481-8865-9_20

Naur, P., 1974. Concise survey of computer methods, 1st edition. ed. Petrocelli Books, New York.

OECD, 2016. Innovating Education and Educating for Innovation The Power of Digital Technologies and Skills [WWW Document]. OECD ILibrary. URL http://www.keepeek.com/Digital-Asset-Management/oecd/education/innovating-education-and-educating-for-innovation_9789264265097-en (accessed 7.9.17).

Ostrow, K.S., Heffernan, N.T., 2016. Studying Learning at Scale with the ASSISTments TestBed, in: Proceedings of the Third (2016) ACM Conference on Learning @ Scale, L@S '16. ACM, New York, NY, USA, pp. 333–334. doi:10.1145/2876034.2893404

Pardo, A., Siemens, G., 2014. Ethical and privacy principles for learning analytics. Br. J. Educ. Technol. 45, 438–450. doi:10.1111/bjet.12152

Pena-Ayala, A., 2014. Educational data mining: A survey and a data mining-based analysis of recent works. Expert Syst. Appl. 41, 1432–1462. doi:10.1016/j.eswa.2013.08.042

Rienties, B., Toetenel, L., Bryan, A., 2015. "Scaling up" learning design: impact of learning design activities on LMS behavior and performance, in: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15. ACM Press, pp. 315–319. doi:10.1145/2723576.2723600

Romero, C., Ventura, S., 2013. Data mining in education. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 3, 12–27. doi:10.1002/widm.1075

Roschelle, J., Feng, M., Murphy, R.F., Mason, C.A., 2016. Online Mathematics Homework Increases Student Achievement. AERA Open 2, 2332858416673968. doi:10.1177/2332858416673968

Scheffel, M., Drachsler, H., Stoyanov, S., Specht, M., 2014. Quality Indicators for Learning Analytics. J. Educ. Technol. Soc. 17, 117.

Sergis, S., Sampson, D.G., 2017. Teaching and Learning Analytics to Support Teacher Inquiry: A Systematic Literature Review, in: Peña-Ayala, A. (Ed.), Learning Analytics: Fundaments, Applications, and Trends, Studies in Systems, Decision and Control. Springer International Publishing, pp. 25–63. doi:10.1007/978-3-319-52977-6_2

Shute, V.J., Rahimi, S., 2017. Review of computer-based assessment for learning in elementary and secondary education. J. Comput. Assist. Learn. 33, 1. doi:10.1111/jcal.12172

Siemens, G., 2012. Learning analytics: envisioning a research discipline and a domain of practice, in: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, pp. 4–8.

Siemens, G., Gasevic, D., Haythornwaite, C., Dawson, S., Shum, S.B., Ferguson, R., Duval, E., Verbert, K., Baker, R.S.J. d, 2011. Open Learning Analytics: an integrated & modularized platform.

Silva, C., Fonseca, J., 2017. Educational Data Mining: A Literature Review, in: Europe and MENA Cooperation Advances in Information and Communication Technologies. Springer, Cham, pp. 87–94. doi:10.1007/978-3-319-46568-5_9

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489. doi:10.1038/nature16961

Slavuj, V., Meštrović, A., Kovačić, B., 2017. Adaptivity in educational systems for language learning: a review. Comput. Assist. Lang. Learn. 30, 64–90. doi:10.1080/09588221.2016.1242502

Smith, F.J., 2006. Data science as an academic discipline. Data Sci. J. 5, 163–164.

Steenbergen-Hu, S., Cooper, H., 2013. A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. J. Educ. Psychol., Advanced Learning Technologies 105, 970–987. doi:10.1037/a0032447

Teichman, A., Thrun, S., 2013. Group induction, in: Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference On. IEEE, pp. 2757–2763.

Toetenel, L., Rienties, B., 2016. Analysing 157 learning designs using learning analytic approaches as a means to evaluate the impact of pedagogical decision making. Br. J. Educ. Technol. 47, 981–992. doi:10.1111/bjet.12423

Wiliam, D., 2010. An integrative summary of the research literature and implications for a new theory of formative assessment. Handb. Form. Assess. 18–40.

Wilson, K., Zack, N., 2015. The Knewton Platform. Gen.-Purp. Adapt. Learn. Infrastruct. Knewton.

Zeiler, M.D., Fergus, R., 2013. Visualizing and Understanding Convolutional Networks. ArXiv13112901 Cs.

Glossary of Key Terms and Acronyms

Term	Definition
Algorithm	A step by step procedure for solving a problem.
Application Programming Interface (API)	Clearly defined methods to enable communication between different software systems.
Big Data	Extremely large data sets.
Classification	Assigning an item or person of interest to a predefined set of categories.
Collaborative Filtering	Automatic predictions about the interests of users based on similar users.
Computer Based Assessment for Learning (CBAfL)	Learners take quizzes and tests on a computer instead of pen and paper which can give detailed feedback to the learner and/or the teacher.
Content Analysis	A method that identifies patterns in large collections of documents.
Data Science	A field about extracting knowledge from data
DELICATE checklist	A set of guidelines that help with the implementation of Learning Analytics.
Domain Based	Knowledge particular to a specific field.
Educational Data Mining (EDM)	Exploring data from an educational setting using a range of statistical or machine learning techniques.
Educational Data Science	A field about extracting knowledge from educational data.
Experience API (xAPI)	Software that allows independent systems to communicate and keep track of any learning experiences.
IMS Caliper/Sensor API	A standard way of describing and gathering learning metrics across different software systems.
Learning Analytics (LA)	The collection, analysis and reporting of data in learning.
Learning Models	A computational model that can be used to describe and understand the behaviour of a learner within a particular learning system or situation.
Learning Record Store (LRS)	A system for storing data about learners used by xAPI and IMS Caliper API.
Machine Learning (ML)	A field in computer science, which deals with enabling computers to learn using existing sets of data without having been explicitly programmed to understand it.
Neural Networks	A computer system which mimics the human brain by using artificial neurons linked together into a network and then trained to recognise certain patterns.
Recommendation Systems	Systems that recommend particular activities or content based on the prior performance of the learner.
Reinforcement Learning	A form of machine learning where the system determines the ideal behaviour by receiving reward feedback when its action is evaluated as being better than previous choices.
Structured Data	Any data that is based on some model, which defines its contents.
Supervised Learning	Represents input patterns as part of a statistical structure with explicit target outputs determined by the designers of the system.

Term	Definition
Unstructured Data	Data with no pre-defined model to explain it.
Unsupervised Learning	Represents input patterns as part of a statistical structure but it has no explicit target outputs specified by the designer of the system.

Technical Appendix

Appendix A- Systematic Review Process

The evidence for this review was gathered using a rapid review protocol that involved a review of surveys, literature reviews, systematic reviews and meta-analyses from the last 5 years combined with a search for new studies of specific techniques, tools and applications in primary and secondary contexts from the last 3 years. Gough, Sandy & James suggest taking this approach where time is limited and it is therefore not possible to conduct a more exhaustive search (Gough et al., 2013, p. 26)

In the first phase Glasgow Universities instance of the Summon 2.0 search engine was used with the search string

(("learning analytics") OR ("educational data mining") OR ("educational data science")) AND ((summary) OR (review) OR (survey) OR (meta-analysis))

and set to search the whole Summon index but only include scholarly & peer reviewed articles with full text online in English and with the content type- journal article, conference proceeding or book chapter. This search string was chosen after an iterative process of experimentation to identify suitable keywords that would only return results for data science and machine learning applied in an educational context.

This initial search yielded 835 relevant results but once duplicates were identified and merged there were 770 unique documents left. This was augmented by a second search with the same search string on Google Scholar in order of relevance to identify additional grey literature and studies missed in the original search. This yielded 14,100 results with the top 1000 reviewed and an additional 158 items identified and added to give 928 unique results. During the review of results of the second search the last two pages of results reviewed contained nothing of relevance in terms of suitable studies for possible inclusion.

The second phase involved looking at the article titles to remove short editorials, calls for papers, acknowledgements and content that was clearly not strongly connected to either learning analytics, educational data science, data science or machine learning which left 535 documents.

In the third phase, the abstracts for these documents were examined by two of the researchers and for articles about the application of data science and machine learning the following inclusion and exclusion criteria were used

Include any documents that were

- applications of data science or machine learning in a primary or secondary education context
- systematic reviews which also included impact on learners or teachers
- analyses of overall trends within EDM and LA research

Exclude any documents that were

- purely focused on higher education
- theoretical in nature that just compared or described machine learning techniques.

- educational technology papers with no mention of learning analytics, educational data mining, data science or machine learning.

For documents focused on exploring challenges in applying EDM or LA more widely

Include any documents that were

- successful approaches that have helped facilitate wider scale adoption of techniques and technologies.

Exclude any documents that were

- solely focused on barriers and difficulties within the context of higher or further education.
- purely speculative difficulties that didn't reference any previous examples or relevant legislation.

This reduced the number of documents to 258 for the final selection process which involved a more careful examination of the content of the document itself. The previous inclusion and exclusion criteria were used again with the addition of the following:

Include any relevant documents that were

- authored by one of the top researchers in the field of learning analytics or educational data mining as identified by the number of unique researchers who cited their work- provided by Google Scholar.

Exclude any documents that were

- overly descriptive and/or mostly repeating large sections of original sources with little additional insight.
- really an in-depth case study of one technology or approach with a very short review or survey at the beginning.
- published in a poor-quality journal that either had no impact factor or extremely low impact factor <= 0.1 as judged using the SCImago index. Some recent open source journals do not currently have an impact factor but were assessed by the range of researchers who published work and how relevant it was to the purpose of this review.

This left 50 documents for the initial analysis and synthesis to help address the following research questions.

1. What definitions are given for key terms such as data science, machine learning, educational data mining and learning analytics?
2. Which types of educational problems can educational data mining and learning analytics be most effectively applied to?
3. Which aspects of education has educational data mining and learning analytics already been applied to?
4. Which problems within education has new technology been developed to address?
5. What has been the impact of the machine learning technology on teachers or learners?
6. What is the timescale for being able to apply specific technologies on a wider scale?
7. What implications does machine learning have for educational research?
8. What are the barriers that could prevent wider scale use of this technology?
9. Which approaches have been successfully used to enable the wide scale adoption of a machine learning application?

As part of this process the primary research source was identified for any points made and then added to the reference collection for this review. The last three rounds of European research funding (FP6, FP7 and Horizon 2020) listed on the CORDIS website have been examined to identify relevant work carried out partly or fully in a primary or secondary school context.

The last three years of the following journals and conference proceedings were also manually examined to identify more recent research in specific areas conducted in the last 3 years resulting in the addition of another 53 documents.

Journal/Conference	**Web Page**
Journal of Educational Data Mining	http://www.educationaldatamining.org/JEDM/index.php/JEDM
International Conference on Educational Data Mining	http://www.educationaldatamining.org/proceedings
Journal of Learning Analytics	http://learning-analytics.info/
International Conference on Learning Analytics and Knowledge	https://solaresearch.org/events/lak/
International Journal on Artificial Intelligence in Education	http://iaied.org/journal/
Artificial Intelligence in Education Conference	http://iaied.org/conf/1/
Journal of the Learning Sciences	http://www.tandfonline.com/toc/hlns20/.VCsyjGeSwrU
International Journal of Computer-Supported Collaborative Learning	http://ijcscl.org
Computers in Human Behavior	https://www.journals.elsevier.com/computers-in-human-behavior

Journal/Conference	Web Page
Computers & Education	https://www.journals.elsevier.com/computers-and-education/
Journal of Educational Technology & Society	http://www.ifets.info/
IEEE Global Engineering Education Conference (EDUCON)	http://educon-conference.org/current/
ACM Conference on Learning @ Scale	http://learningatscale.acm.org/
Journal of Computer Assisted Learning	http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-2729
British Journal of Educational Technology	http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1467-8535
Artificial Intelligence Review	https://link.springer.com/journal/10462
Expert Systems with Applications	https://www.journals.elsevier.com/expert-systems-with-applications/

Appendix B- Highly Cited Researchers in Educational Data Science Fields

These lists of researchers in alphabetical order are generated from the researchers identified as working within a particular subfield and how highly cited they are in general.

Educational Data Mining		
Forename	**Surname**	**Other Education Related Areas of Expertise**
*Ryan	Baker	Learning Analytics, Engagement, Affect, Learning
Joseph	Beck	Student Modeling, Adaptivity, Causal Relations
Nabil	Belacel	Supervised and Unsupervised Learning
Marie	Bienkowski	Learning Analytics, Open Educational Resources
Christopher A	Brooks	E-Learning, HCI, AI, Learning Analytics
Richard	Cox	Cognitive Science, Diagrammatic Reasoning, Technology Enhanced Learning
Michel	Desmarais	AI, Student Modelling, HCI
Cyril	Goutte	Machine Learning, Natural Language Processing
*Neil	Heffernan	AI, User Modeling, Intelligent Tutoring Systems
Richard A.	Levine	Computational Statistics, Bayesian Analysis
Collin	Lynch	AI, AI in Education, Electronic Voting
Ma.	Mercedes Rodrigo	AI in Education, Affective Computing
Agathe	Merceron	Information Visualisation
Julià	Minguillón	E-Learning, Learning Analytics, Visualisation
*Jack	Mostow	AI in Education, Intelligent Tutors, Speech Understanding
Sreerama KV	Murthy	Machine Learning, Image Understanding, Technology Enabled Learning
*Hiroaki	Ogata	Learning Analytics, CSCL, Technology Enhanced Learning
Saurabh	Pal	
Zachary A.	Pardos	Learning Analytics, Machine Learning
*Radek	Pelánek	AI in Education, Learning Analytics, Intelligent Tutoring Systems
Cristobal	Romero	
Richard	Scheines	Causation, Causal Discovery and Latent Variable Models
*Daniel T.	Seaton	Learning Analytics, Physics Education Research, Computational Physics
*Kalina	Yacef	AI in Education, HCI, Intelligent Tutoring Systems
*Stephen JH	Yang	Mobile learning, AI in education, Learning Analytics, MOOCs

Learning Analytics		
Forename	**Surname**	**Other Education Related Areas of Expertise**
*Ryan	Baker	Educational Data Mining, Engagement, Affect
Simon J.	Buckingham Shum	Argument Mapping, Collective Intelligence, HCI
Susan	Bull	AI in Education, Technology Enhanced Learning, User Model, Personalisation
Shane	Dawson	Social Network Analysis, Technology Enhanced Learning, Learning Design
Sara	de Freitas	Educational Technology, Pedagogy, Game Science, Learner Engagement
Erik	Duval	Technology Enhanced Learning, HCI, Quantified Self
Martin	Ebner	E-Learning, Open Educational Resources, MOOC, Technology Enhanced Learning
Baltasar	Fernández Manjón	E-Learning, Serious Games, Medical Simulation, Games Based Learning
Dragan	Gasevic	Self-regulated Learning, Technology Enhanced Learning, Collaborative Learning, Learning Technologies
Sabine	Graf	Technology Enhanced Learning, Adaptive Learning Systems, User Modeling, Mobile Learning
George	Karypis	Recommender Systems
Judy	Kay	HCI, AI in Education, Technology Enhanced Learning
Gregor	Kennedy	Educational Technology, eLearning, Online Learning
Ken	Koedinger	Learning Technologies, AI in Education, Learning Sciences, HCI
Min	Liu	Multimedia, Problem/Project-based, Games Based Learning, Emerging Technologies
Rose	Luckin	Learning Sciences
Timothy	McKay	
*Hiroaki	Ogata	Mobile and Ubiquitous Learning, Educational Data Mining, CSCL, Technology Enhanced Learning
Abelardo	Pardo	Learning Technologies, Technology Enhanced Learning, Learning Design, Educational Technology
Demetrios	Sampson	Learning Technologies, Technology Enhanced Learning, eLearning, Educational Decision Support Systems
George	Siemens	Social Networks, Collective Intelligence, Networked Learning
*Daniel	Suthers	CSCL, Technology Enhanced Learning, Social Informatics
Karen	Swan	Online Learning, Data Literacy, Programming & Problem Solving, Community of Inquiry Framework
Vassilios	Verykios	Privacy Preserving Record Linkage
David	Wiley	Open Educational Resources, Learning Objects
*Stephen JH	Yang	Mobile Learning, AI in Education, Educational Data Mining

Intelligent Tutoring Systems		
Forename	**Surname**	**Other Education Related Areas of Expertise**
Vincent	Aleven	Learning Science and Technologies, Educational Games
Ivon	Arroyo	Learning Sciences, Mathematics Education, Games for Learning
Cristina	Conati	User Adaptive Interaction
Claude	Frasson	Brain Computing
Janice	Gobert	Science Learning and Assessment, Epistemology
Peter	Hastings	Serious Games, Natural Language Processing
*Neil	Hefferman	Educational Data Mining, User Modeling, Cognition of Mathematics Learning
Tsukasa	Hirashima	Learning Engineering, Technology Enhanced Learning, Educational Technology, STEM Learning
Xiangen	Hu	Cognitive Psychology, Semantic Analysis
Tanner	Jackson	Educational Games
Amruth	Kumar	Model-Based Reasoning, Computer Science Education
James	Lester	AI in Education, Educational Technology, Learning Technologies
Diane	Litman	Natural Language Processing, AI in Education
Bruce M.	McLaren	Computers and Education, Educational Games, CSCL
Riichiro	Mizoguchi	Applied Ontology, AI in Education, Semantic Web
Bradford	Mott	Games Based Learning Environments, Interactive Narrative, Computer Games
Tom	Murray	Educational Technology, Deliberative Skills, Knowledge Engineering, Inquiry Learning
John C.	Nesbit	Educational Psychology, Self-regulated Learning, Multimedia Learning
Roger	Nkambou	Knowledge-Based Systems, Cognitive Agents, Computer Education
Andrew	Olney	Vector Space Models, Dialogue Systems, Unsupervised Grammar Induction
Beverly	Park Woolf	AI in Education, Educational Technology
*Radek	Pelanek	Educational Data Mining, AI in Education, Learning Analytics
Sergey	Sosnovsky	E-Learning, Adaptive Technologies, User Modeling
Kurt	VanLehn	
Rosa Maria	Vicari	Informatics in Education

Computer Supported Collaborative Learning		
Forename	**Surname**	**Other Education Related Areas of Expertise**
Jerry	Andriessen	Interactive Technology, Argumentation
Juan I.	Asensio-Perez	Educational Technology, Technology Enhanced Learning, Learning Design
Michael	Baker	Collaborative Learning, CSCW
Bram	De Wever	Technology Enhanced Learning, Asynchronous Discussion Groups
Pierre	Dillenbourg	Learning Technologies, HCI, MOOCs,
Yannis	Dimitriadis	Educational Technology, Computer Science Technology, Technology Enhanced Learning, Learning Design
Andrea	Forte	HCI, Social Computing, CSCW
Hugo	Fuks	Wearables
Cindy	Hmelo-Silver	Cognition and Instruction, Problem Based Learning
Patrick	Jermann	CSCW, Eye-Tracking, Pair Programming, Tangible Computing
Chris	Jones	Networked Learning, Technology Enhanced Learning, Computer Assisted Learning
Gellof	Kanselaar	Psychology, ICT, Learning
Joachim	Kimmerle	Educational Psychology, Knowledge Media, Medical Education
Karel	Kreijns	Open Educational Resources, MOOCs
*Alejandra	Martinez Mones	Learning Analytics, Technology Enhanced Learning, Social Network Analysis
Miguel	Nussbaum	Collaborative Learning, Educational Technology, Technology Enhanced Learning, Instructional Design
*Hiroaki	Ogata	Mobile & Ubiquitous Learning, Educational Data Mining, Learning Analytics, Technology Enhanced Learning
Daniel	Suthers	Technology Enhanced Learning, Social Informatics
Deborah	Tatar	HCI, CSCW, Microcoordination
Stephanie D.	Teasley	HCI, CSCW
Michael	Twidale	HCI, CSCW, Informal Learning of Technology
Hao-Chaun	Wang	HCI, Social Computing, CSCW
Rupert	Wegerif	Dialogic Education, Teaching Thinking, Educational Theory
Armin	Weinberger	Collaboration Scripts, Knowledge Convergence, Intercultural Learning, Argumentative Knowledge Construction
Jianwei	Zhang	Collaborative Learning, Knowledge Building, Learning Sciences, Educational Technology

Appendix C- Research Projects and Organisation Websites for Educational Data Science

Project Name	**Description**
AFEL Project Analytics for Everyday Learning http://afel-project.eu/	The goal of AFEL (Analytics for Everyday Learning) is to develop methods and tools to understand informal/collective learning as it surfaces implicitly in online social environments. The aim of AFEL is to devise the tools for exploiting learning analytics on such learning activities, in relation to cognitive models of learning and collaboration that are necessary to the understanding of loosely defined learning processes in online social environments.
ASSiSTments https://www.assistments.org/	Assistments is a tool for assigning work that supports greater student learning through the use of immediate feedback. At its core, the site is an online form of homework with the added perk of help embedded in each question. Teachers create sections of classes in which their students enrol, and then assign work to a specific section. Once this happens, students will be able to access the assignment and complete it at their own pace, taking breaks if needed.
K4A Project Knowledge for All http://www.k4all.org/	Knowledge 4 All Foundation Ltd (K4A) is a distributed institute based in London (UK), with two main streams of activities, on one side pioneering Machine Learning methods of pattern analysis, statistical modelling, and computational learning and on the other transform these into technologies for large scale applications in Open Education. It is therefore an advocate of Artificial Intelligence and Big data in Open Education.
Lea's Box Project http://www.leas-box.eu	LEA's BOX is a learning analytics toolbox that is intended to enable educators to perform competence-centred, multi-source learning analytics
LinkedUp Project https://linkedup-project.eu/	The LinkedUp Project is a FP7 (European Framework Programme) Support Action which pushes forward the exploitation and adoption of public, open data available on the Web, in particular by educational organisations and institutions.

LTfLL Project http://www.ltfll-project.org/	The two baseline objectives of 'Language Technologies for Lifelong Learning' project (LTfLL) are: (a) helping people learn, and, (b) helping tutors/teachers to support learners. LTfLL is directed towards text-based artefacts and uses several different language technologies to analyse them and to give feedback about them back to the users.
Newton Project Networked labs for training in sciences and technologies http://www.newtonproject.eu/	NEWTON develop, integrate and disseminate innovative technology enhanced learning (TEL) methods and tools, to create or inter-connect existing state-of the art teaching labs and to build a pan-European learning network platform that supports fast dissemination of learning content to a wide audience in a ubiquitous manner. NEWTON focuses on employing novel technologies to increase learner quality of experience, improve learning process and increase learning outcome.
NEXT-TELL Project Next Generation Teaching, Education and Learning for Life http://next-tell.eu/	NEXT-TELL's foremost goal is to provide an innovation platform, used by teachers to continuously and collaboratively innovate ICT-enhanced formative classroom assessment.
PELARS Project http://www.learningmaking.eu/	Designed and developed real technologies (kits, electronics systems, and sensor systems in the classroom) that help teachers and learners understand what is happening when people do Science Technology Engineering and MATH (STEM).
PRIMES Project https://primes-project.eu/	PRIMES aims to produce a novel e-learning platform, combining the wide reach and popularity of MOOCs with advanced artificial intelligence and data summarisation algorithms, to provide a truly personalised learning experience, fitted to the needs and characteristics of teenager (secondary school) students.
Siette Project https://www.siette.org/	SIETTE is a domain-independent assessment environment that supports different assessment methods — including classical test theory, item response theory, and computer adaptive testing and integrates them with multidimensional student models used by intelligent educational systems. Teachers can use an authoring tool to create large item pools of different types of questions, including multiple choice, open answer, generative questions, and complex tasks.

STELA Project http://stela-project.eu/	The main goal of the project is to enhance a successful transition from secondary to higher education by means of learning analytics. To this end the project will develop, test, and assess a learning analytics approach that focuses on providing formative and summative feedback to students in the transition.
TeSLA Project http://tesla-project.eu/	The TeSLA project provides to educational institutions, an adaptive trust e-assessment system for assuring e-assessment processes in online and blended environments. It will support both continuous and final assessment to improve the trust level across students, teachers and institutions. The system will be developed taking into account quality assurance agencies in education, privacy and ethical issues and educational and technological requirements throughout Europe.
Organisation Name	**Description**
CORDIS Community Research and Development Information Service http://cordis.europa.eu/projects/home_en.html	CORDIS is the Community Research and Development Information Service. It is the European Commission's primary public repository and portal to disseminate information on all EU-funded research projects and their results in the broadest sense
International Educational Data Mining Society http://www.educationaldatamining.org/	The International Educational Data Mining Society was founded in July 2011, by the International Working Group on Educational Data Mining. The Society works towards increasing opportunities for participation, input from the broader EDM community and maintaining scientific quality while setting up an organization that can function smoothly, indefinitely.
Jisc Effective Learning Analytics Project https://www.jisc.ac.uk/rd/projects/effective-learning-analytics	The Effective Learning Analytics challenge is about using data and analytics to support students; improving satisfaction, retention and graduation rates.
LACE Evidence Hub http://evidence.laceproject.eu/	The LACE Evidence Hub can help you make evidence–based decisions about learning and teaching, whether you are a teacher, a manager, a researcher or a policymaker.

Learning Analytics Community Europe Project http://www.laceproject.eu/	LACE partners are passionate about the opportunities afforded by current and future views of learning analytics (LA) and educational data mining (EDM). The project's aim is to integrate communities working on LA and EDM from schools, workplace and universities by sharing effective solutions to real problems.
Society for Learning Analytics Research https://solaresearch.org/	The Society for Learning Analytics Research (SoLAR) is an inter-disciplinary network of leading international researchers who are exploring the role and impact of analytics on teaching, learning, training and development. SoLAR has been active in organizing the International Conference on Learning Analytics & Knowledge (LAK) and the Learning Analytics Summer Institute (LASI), launching multiple initiatives to support collaborative and open research around learning analytics, promoting the publication and dissemination of learning analytics research, and advising and consulting with state, provincial, and national governments.
The International Intelligence in Education Society (IAIED) http://iaied.org/about/	AIED is an interdisciplinary community at the frontiers of the fields of computer science, education and psychology. It promotes rigorous research and development of interactive and adaptive learning environments for learners of all ages, across all domains. The society brings together a community of members in the field through the organization of Conferences, a Journal, and other activities of interest.

Appendix D- Existing Tools, Applications and Products

Products for Learners and Educators	
ALEKS https://www.aleks.com/	ALEKS Corporation is a leader in the creation of Web-based, artificially intelligent, educational software. ALEKS assessment and learning technologies were originally developed by a team of cognitive scientists and software engineers at the University of California, Irvine, with major funding from the National Science Foundation. ALEKS is founded on ground-breaking research in mathematical cognitive science known as Knowledge Space Theory. Through adaptive questioning, ALEKS accurately assesses a student's knowledge state and then delivers targeted instruction on the exact topics the student is most ready to learn.
Carnegie Learninglearn http://www.carnegielearning.com	Carnegie Learning is focused on transforming the math classroom, and putting better math learning within reach for both teachers and students.
CogBooks https://www.cogbooks.com/	CogBooks is committed to transforming the way teachers teach and students learn, by applying science-based methods to education. Through unique advanced adaptive learning technologies, they're changing the way secondary and further education students learn in the US, UK and worldwide.
Connect 2 https://www.mheducation.co.uk/connect2	Connect2 is a Learning Science company which means that they research and study how learning happens. Based on this, methods to make the learning process easier and more effective are developed.
DreamBox Learning http://www.dreambox.com/	DreamBox Learning is an online software provider that focuses on mathematics education at the elementary and middle school level.
Fuse Schools https://www.fuseschool.org/	FuseSchool is a global Open Education charity, brought into life by the leading Social Learning Organisation Fuse Universal. Their mission is to provide great education to millions of learners who need it most - so that they can improve their lives and communities.

Gooru http://www.gooru.org/	Gooru's learning navigator provides real-time data to help students understand where they are and how to reach their goals. Gooru blends powerful assessments and analytics with a catalog of learning resources to help teachers track student progress and understanding and meet the needs of every student.
Knewton https://www.knewton.com/	Knewton is an adaptive learning company that has developed a platform to personalize educational content as well as has developed courseware for higher education concentrated in the fields of science, technology, engineering, and mathematics.
Realizeit http://realizeitlearning.com/	Realizeit wants to provide every learner the opportunity to achieve his or her full learning potential. They make that happen by empowering institutions to significantly improve student motivation, engagement and achievement through the deployment of a real-time evidence-based learning system that enables adaptive learning, adaptive teaching, institutional effectiveness and more.
SumDog https://www.sumdog.com/	Sumdog's free educational games motivate students to practise maths, reading and writing - while you have full control over what they learn.
Cognitive Computing Platforms	
Amazon AI https://aws.amazon.com/amazon-ai/	Amazon AI services bring natural language understanding (NLU), automatic speech recognition (ASR), visual search and image recognition, text-to-speech (TTS), and machine learning (ML) technologies within the reach of every developer. Based on the same proven, highly scalable products and services built by the thousands of machine learning experts across Amazon.
Coseer https://coseer.com/	With Coseer, tasks that took days take milliseconds. You make faster, more informed decisions and can expand the scope of your business activity by an order of magnitude. Coseer is configured to emulate your best performers, and for the nuances specific to your workflow. Then trained to 95-98% accuracy, so you can depend on it. Your time and resources are precious. Coseer requires no pre-labelling, tagging or data-cleaning, and comes to life within 4-12 weeks.

Google Cloud Machine Learning Engine https://cloud.google.com/ml-engine/	Google Cloud Machine Learning Engine is a managed service that enables you to easily build machine learning models that work on any type of data, of any size. Create your model with the powerful TensorFlow framework that powers many Google products, from Google Photos to Google Cloud Speech. Build models of any size with our managed scalable infrastructure.
Microsoft Azure Cognitive Services https://azure.microsoft.com/en-gb/services/cognitive-services/	Microsoft Cognitive Services let you build apps with powerful algorithms using just a few lines of code. They work across devices and platforms such as iOS, Android and Windows, keep improving and are easy to set up.
Watson Developer Cloud https://www.ibm.com/watson/developercloud/	Watson developer cloud is a range of smart services that provide help processing and understanding speech and visual input as well as making sense of textual information from a range of unstructured and structured sources.

Appendix E- Complete List of Educational Data Science Documents Analysed

Adams Becker, S., Freeman, A., Gieseinger Hall, C., Cummins, M., Yuhnke, B., 2016. NMC/CoSN Horizon Report 2016 K-12 Edition. The New Media Consortium.

Agudo-Peregrina, Á.F., Iglesias-Pradas, S., Conde-González, M.Á., Hernández-García, Á., 2014. Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. Computers in Human Behavior 31, 542–550. doi:10.1016/j.chb.2013.05.031

Algarni, A., 2016. Data Mining in Education. International Journal of Advanced Computer Science and Applications 7. doi:10.14569/IJACSA.2016.070659

Almosallam, E.A., Ouertani, H.C., 2014. Learning Analytics: definitions, applications and related fields, in: Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013). Springer, Singapore, pp. 721–730. doi:10.1007/978-981-4585-18-7_81

Al-Razgan, M., Al-Khalifa, A.S., Al-Khalifa, H.S., 2014. Educational Data Mining: A Systematic Review of the Published Literature 2006-2013, in: Herawan, T., Deris, M.M., Abawajy, J. (Eds.), Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013). Springer Singapore, Singapore, pp. 711–719. doi:10.1007/978-981-4585-18-7_80

Amara, S., Macedo, J., Bendella, F., Santos, A., 2016. Group Formation in Mobile Computer Supported Collaborative Learning Contexts: A Systematic Literature Review. Journal of Educational Technology & Society 19, 258.

Apereo Foundation, 2016. Apereo Open Analytics An open platform for optimising learning.

Baker, R.S., Inventado, P.S., 2014. Educational Data Mining and Learning Analytics, in: Larusson, J.A., White, B. (Eds.), Learning Analytics. Springer New York, pp. 61–75. doi:10.1007/978-1-4614-3305-7_4

Berland, M., Davis, D., Smith, C.P., 2015. AMOEBA: Designing for collaboration in computer science classrooms through live learning analytics. Intern. J. Comput.-Support. Collab. Learn 10, 425–447. doi:10.1007/s11412-015-9217-z

Bienkowski, M., Feng, M., Means, B., 2012. Enhancing teaching and learning through educational data mining and learning analytics: An issue brief, in: Proceedings of Conference on Advanced Technology for Education.

Blanco, Á. del, Serrano, Á., Freire, M., Martínez-Ortiz, I., Fernández-Manjón, B., 2013. E-Learning standards and learning analytics. Can data collection be improved by using standard data models?, in: 2013 IEEE Global Engineering Education Conference (EDUCON). Presented at the 2013 IEEE Global Engineering Education Conference (EDUCON), pp. 1255–1261. doi:10.1109/EduCon.2013.6530268

Blikstein, P., Worsley, M., 2016. Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. Journal of Learning Analytics 3, 220–238. doi:10.18608/jla.2016.32.11

Bousbia, N., Belamri, I., 2014. Which Contribution Does EDM Provide to Computer-Based Learning Environments?, in: Peña-Ayala, A. (Ed.), Educational Data Mining, Studies in Computational Intelligence. Springer International Publishing, pp. 3–28. doi:10.1007/978-3-319-02738-8_1

Buckingham Shum, S., Hawksey, M., Baker, R.S., Jeffery, N., Behrens, J.T., Pea, R., 2013. Educational data scientists: a scarce breed, in: Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM, pp. 278–281.

Chatti, M.A., Dyckhoff, A.L., Schroeder, U., Thus, H., 2012. A Reference Model for Learning Analytics. International Journal of Technology Enhanced Learning (IJTEL).

Clow, D., 2013. An overview of learning analytics. Teaching in Higher Education 18, 683–695. doi:10.1080/13562517.2013.827653

Clow, D., Cross, S., Ferguson, R., Rienties, B., 2014. Evidence Hub Review. The Open University, UK, for the LACE Project (http://www.laceproject.eu).

Cope, B., Kalantzis, M., 2016. Big Data Comes to School: Implications for Learning, Assessment, and Research. AERA Open 2, 2332858416641907. doi:10.1177/2332858416641907

Dai, Y., Asano, Y., Yoshikawa, M., 2016. Course Content Analysis: An Initiative Step toward Learning Object Recommendation Systems for MOOC Learners, in: Proceedings of the 9th International Conference on Educational Data Mining.

Daley, S.G., Hillaire, G., Sutherland, L.M., 2016. Beyond performance data: Improving student help seeking by collecting and displaying influential data in an online middle-school science curriculum. Br J Educ Technol 47, 121–134. doi:10.1111/bjet.12221

Dede, C.J., 2016. Next steps for “Big Data” in education: Utilizing data-intensive research.

Doroudi, S., Holstein, K., Aleven, V., Brunskill, E., 2016. Sequence Matters, But How Exactly? A Method for Evaluating Activity Sequences from Data., in: EDM. pp. 70–77.

Drachsler, H., Greller, W., 2016. Privacy and Analytics: It’s a DELICATE Issue a Checklist for Trusted Learning Analytics, in: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK ’16. ACM, New York, NY, USA, pp. 89–98. doi:10.1145/2883851.2883893

Drachsler, H., Greller, W., 2012. The Pulse of Learning Analytics Understandings and Expectations from the Stakeholders, in: Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge, LAK ’12. ACM, New York, NY, USA, pp. 120–129. doi:10.1145/2330601.2330634

Ebner, M., Schön, M., 2013. Why learning analytics in primary education matters. Bulletin of the Technical Committee on Learning Technology 15, 14–17.

Essa, A., 2016. A possible future for next generation adaptive learning systems. Smart Learning Environments 3, 1–24. doi:10.1186/s40561-016-0038-y

Evidence paper RF DC.pdf, n.d.

Feild, J.L., Lewkow, N., Zimmerman, N., Riedesel, M., Essa, A., 2016. A Scalable Learning Analytics Platform for Automated Writing Feedback., in: EDM. pp. 688–693.

Ferguson, R., 2012. Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning 4, 304. doi:10.1504/IJTEL.2012.051816

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., Vuorikari, R., 2016. Research Evidence on the Use of Learning Analytics: Implications for Education Policy.

Ferguson, R., Clow, D., 2017. Where is the evidence?: a call to action for learning analytics, in: Proceedings of the Seventh International Learning Analytics & Knowledge Conference LAK '17. ACM Press, pp. 56–65. doi:10.1145/3027385.3027396

Fyfe, E.R., 2016. Providing feedback on computer-based algebra homework in middle-school classrooms. Computers in Human Behaviour 63, 568–574.

Ganesh, S.H., Christy, A.J., 2015. Applications of Educational Data Mining: A survey, in: 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS). Presented at the 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1–6. doi:10.1109/ICIIECS.2015.7192945

Gibson, D.C., Webb, M.E., 2015. Data science in educational assessment. Education and Information Technologies. doi:10.1007/s10639-015-9411-7

Goggins, S.P., Xing, W., Chen, X., Chen, B., Wadholm, B., 2015. Learning Analytics at" Small" Scale: Exploring a Complexity-Grounded Model for Assessment Automation. Journal of Universal Computer Science 21, 66–92.

Griffiths, D., Drachsler, H., Kickmeier-Rust, M., Steiner, C., Hoel, T., Greller, W., 2016. Is Privacy a Show-stopper for Learning Analytics? A Review of Current Issues and Solutions. Learning Analytics Community Exchange 6.

Hansen, C.J., Wasson, B., 2016. Teacher Inquiry into Student Learning: - The TISL Heart Model and Method for use in Teachers' Professional Development. Nordic Journal of Digital Literacy 11, 24–49. doi:10.18261/issn.1891-943x-2016-01-02

Howlin, C., Lynch, D., 2014. A framework for the delivery of personalized adaptive content, in: 2014 International Conference on Web and Open Access to Learning (ICWOAL). Presented at the 2014 International Conference on Web and Open Access to Learning (ICWOAL), pp. 1–5. doi:10.1109/ICWOAL.2014.7009203

Hu, X., Cheong, C.W.L., Ding, W., Woo, M., 2017. A systematic review of studies on predicting student learning outcomes using learning analytics. ACM Press, pp. 528–529. doi:10.1145/3027385.3029438

Johnson, L., Adams Becker, S., Estrada, V., Freeman, A., Kampylis, P., Vuorikari, R., Punie, Y., 2014. The NMC horizon report Europe: 2014 schools edition., NMC horizon report. Publications Office, Luxembourg.

Khalil, M., Ebner, M., 2016. What is Learning Analytics about? A Survey of Different Methods Used in 2013-2015. Smart Learning Environments.

Khalil, M., Ebner, M., 2015. Learning Analytics: Principles and Constraints. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2015 1326–1336.

Kickmeier-Rust, M., 2017. Tools for Competence-Centred, Multi-Source Learning Analytics Final Report (Final Report).

Kickmeier-Rust, M., Steiner, C.M., Albert, D., 2014. Learning Analytics and Educational Data Mining: An Overview of Recent Techniques. Learning Analytics For and In Serious Games.

Kickmeier-Rust, M.D., Albert, D., 2016. Theory-driven Learning Analytics and Open Learner Modelling: The Teacher's Toolbox of Tomorrow?, in: UMAP (Extended Proceedings).

Klašnja-Milićević, A., Ivanović, M., Nanopoulos, A., 2015. Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions. The Artificial Intelligence Review 44, 571–604. doi:10.1007/s10462-015-9440-z

Koedinger, K.R., Baker, R.Sj., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J., 2010. A data repository for the EDM community: The PSLC DataShop, in: Handbook of Educational Data Mining.

Koedinger, K.R., D'Mello, S., McLaughlin, E.A., Pardos, Z.A., Rosé, C.P., 2015. Data mining and education. WIREs Cogn Sci 6, 333–353. doi:10.1002/wcs.1350

Konya, K., Burgos, D., Fuente Valentín, L. de la, Carrasco, A., 2013. Emerging technologies landscape on education: A review. IJIMAI 2, 55.

Kopeinik, S., Kowald, D., Lex, E., 2016. Which Algorithms Suit Which Learning Environments? A Comparative Study of Recommender Systems in TEL, in: Adaptive and Adaptable Learning. Presented at the European Conference on Technology Enhanced Learning, Springer, Cham, pp. 124–138. doi:10.1007/978-3-319-45153-4_10

Lane, C., Grover, S., Roschelle, J., 2016. AI Applications in Education. The Center for Innovative Research in CyberLearning,[Online]. Available: http://circlcenter. org/ai-applications-education/.| Accessed February.

Li, K.C., Lam, H.K., Lam, S.S., 2015. A Review of Learning Analytics in Educational Research, in: Technology in Education. Technology-Mediated Proactive Learning. Presented at the International Conference on Technology in Education, Springer, Berlin, Heidelberg, pp. 173–184. doi:10.1007/978-3-662-48978-9_17

Liu, S., Kunnan, A., 2016. Investigating the Application of Automated Writing Evaluation to Chinese Undergraduate English Majors. CALICO 33, 71–91. doi:10.1558/cj.v33i1.26380

Luckin, R., Clark, W., Avramides, K., Hunter, J., Oliver, M., 2017. Using teacher inquiry to support technology-enhanced formative assessment: a review of the literature to inform a new method. Interactive Learning Environments 25, 85–97. doi:10.1080/10494820.2015.1121152

Lynch, T., Ghergulescu, I., 2017. Review of Virtual Labs As The Emerging Technologies For Teaching STEM Subjects, in: Proceedings of the 11th International Technology, Education and Development Conference. Presented at the 11th International Technology, Education and Development Conference.

Ma, W., Adesope, O.O., Nesbit, J.C., Liu, Q., 2014. Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology 106, 901–918. doi:10.1037/a0037123

Márquez-Vera, C., Cano, A., Romero, C., Noaman, A.Y.M., 2016. Early dropout prediction using data mining: a case study with high school students. Expert Systems 33, 107–124. doi:10.1111/exsy.12135

Mayer, H., 2015. NEXT-TELL Project Final Report September 2010 - October 2014.

McKenney, S.E., Mor, Y., 2015. Supporting teachers in data-informed educational design. British Journal of Educational Technology 46, 265–279.

Métayer, D.L., 2010. Privacy by Design: A Matter of Choice, in: Data Protection in a Profiled World. Springer, Dordrecht, pp. 323–334. doi:10.1007/978-90-481-8865-9_20

MJRodriguezTriana et al. - 2016 - LA in small-scale teacher-led innovations.pdf, n.d.

Monroy, C., Rangel, V.S., Whitaker, R., 2013. STEMscopes: Contextualizing Learning Analytics in a K-12 Science Curriculum, in: Proceedings of the Third International Conference on Learning Analytics and Knowledge, LAK '13. ACM, New York, NY, USA, pp. 210–219. doi:10.1145/2460296.2460339

Novik, A., Firtova, L., Ali Turker, M., Kickmeier-Rust, M., Bedeck, M., Johnson, M., Ginon, B., Bull, S., 2016. LEA's Box Piloting and Evaluation Report 2.

Ochoa, X., Suthers, D., Verbert, K., Duval, E., 2014. Analysis and Reflections on the Third Learning Analytics and Knowledge Conference (LAK 2013). Journal of Learning Analytics 1, 5–22. doi:10.18608/jla.2014.12.2

OECD, 2016. Innovating Education and Educating for Innovation The Power of Digital Technologies and Skills [WWW Document]. OECD iLibrary. URL http://www.keepeek.com/Digital-Asset-Management/oecd/education/innovating-education-and-educating-for-innovation_9789264265097-en (accessed 7.9.17).

Ostrow, K.S., Heffernan, N.T., 2016. Studying Learning at Scale with the ASSISTments TestBed, in: Proceedings of the Third (2016) ACM Conference on Learning @ Scale, L@S '16. ACM, New York, NY, USA, pp. 333–334. doi:10.1145/2876034.2893404

p528-hu.pdf, n.d.

Papamitsiou, Z., Economides, A.A., 2016. Learning Analytics for Smart Learning Environments: A Meta-Analysis of Empirical Research Results from 2009 to 2015, in: Spector, M.J., Lockee, B.B., Childress, M.D. (Eds.), Learning, Design, and Technology. Springer International Publishing, pp. 1–23. doi:10.1007/978-3-319-17727-4_15-1

paper_169.pdf, n.d.

Pardo, A., Siemens, G., 2014. Ethical and privacy principles for learning analytics. British Journal of Educational Technology 45, 438–450. doi:10.1111/bjet.12152

Pena-Ayala, A., 2014. Educational data mining: A survey and a data mining-based analysis of recent works. Expert Systems with Applications 41, 1432–1462. doi:10.1016/j.eswa.2013.08.042

Persico, D., Pozzi, F., 2015. Informing learning design with learning analytics to improve teacher inquiry. British Journal of Educational Technology 46, 230. doi:10.1111/bjet.12207

Prasad, D., Totaram, R., Usagawa, T., 2016. Development of Open Textbooks Learning Analytics System. International Review of Research in Open and Distance Learning 17.

Reinhardt, W., Meier, C., Drachsler, H., Sloep, P., 2011. Analyzing 5 Years of EC-TEL Proceedings, in: Towards Ubiquitous Learning, Lecture Notes in Computer Science. Presented at the European Conference on Technology Enhanced Learning, Springer, Berlin, Heidelberg, pp. 531–536. doi:10.1007/978-3-642-23985-4_51

Reyes, J.A., 2015. The skinny on big data in education: Learning analytics simplified. TECHTRENDS TECH TRENDS 59, 75–80. doi:10.1007/s11528-015-0842-1

Rienties, B., Toetenel, L., Bryan, A., 2015. "Scaling up" learning design: impact of learning design activities on LMS behavior and performance, in: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15. ACM Press, pp. 315–319. doi:10.1145/2723576.2723600

Rodriguez-Triana, M.J., Martinez-Mones, A., Villagrá-Sobrino, S., 2016. Learning analytics in small-scale teacher-led innovations: Ethical and data privacy issues. Journal of Learning Analytics.

Romero, C., Ventura, S., 2017. Educational data science in massive open online courses. WIREs Data Mining Knowl Discov 7, n/a-n/a. doi:10.1002/widm.1187

Romero, C., Ventura, S., 2013. Data mining in education. WIREs Data Mining Knowl Discov 3, 12–27. doi:10.1002/widm.1075

Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J. d, 2010. Introduction of the handbook of educational data mining, in: Handbook of Educational Data Mining. CRC Press.

Roschelle, J., Feng, M., Murphy, R.F., Mason, C.A., 2016. Online Mathematics Homework Increases Student Achievement. AERA Open 2, 2332858416673968. doi:10.1177/2332858416673968

Santos, J.L., Verbert, K., Klerkx, J., Duval, E., Charleer, S., Ternier, S., 2015. Tracking data in open learning environments. Journal of Universal Computer Science 21, 976–996.

Scheffel, M., Drachsler, H., Stoyanov, S., Specht, M., 2014. Quality Indicators for Learning Analytics. Journal of Educational Technology & Society 17, 117–132.

Sergis, S., Sampson, D.G., 2017. Teaching and Learning Analytics to Support Teacher Inquiry: A Systematic Literature Review, in: Peña-Ayala, A. (Ed.), Learning Analytics: Fundaments, Applications, and Trends, Studies in Systems, Decision and Control. Springer International Publishing, pp. 25–63. doi:10.1007/978-3-319-52977-6_2

Sergis, S., Sampson, D.G., 2016. Towards a Teaching Analytics Tool for Supporting Reflective Educational (re) Design in Inquiry-Based STEM Education, in: 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT). Presented at the 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT), pp. 314–318. doi:10.1109/ICALT.2016.134

Shum, S.B., 2012. Learning Analytics Policy Brief. UNESCO Institute for Information Technologies in Education.

Shute, V.J., Rahimi, S., 2017. Review of computer-based assessment for learning in elementary and secondary education. Journal of Computer Assisted Learning 33, 1. doi:10.1111/jcal.12172

Siemens, G., 2013. Learning Analytics The Emergence of a Discipline. American Behavioral Scientist 57, 1380–1400. doi:10.1177/0002764213498851

Siemens, G., 2012. Learning analytics: envisioning a research discipline and a domain of practice, in: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, pp. 4–8.

Siemens, G., Gasevic, D., Haythornwaite, C., Dawson, S., Shum, S.B., Ferguson, R., Duval, E., Verbert, K., Baker, R.S.J. d, 2011. Open Learning Analytics: an integrated & modularized platform.

Silva, C., Fonseca, J., 2017. Educational Data Mining: A Literature Review, in: Europe and MENA Cooperation Advances in Information and Communication Technologies. Springer, Cham, pp. 87–94. doi:10.1007/978-3-319-46568-5_9

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489. doi:10.1038/nature16961

Sin, K., Muthu, L., 2015. Application of Big Data in Education Data Mining and Learning Analytics – a Literature Review. ICTACT Journal on Soft Computing 5, 1035–1049. doi:10.21917/ijsc.2015.0145

Slater, S., Joksimovic, S., Kovanovic, V., Baker, R.S., 2017. Tools for Educational Data Mining. Journal of Educational and Behavioral Statistics 42, 85. doi:10.3102/1076998616666808

Slavuj, V., Meštrović, A., Kovačić, B., 2017. Adaptivity in educational systems for language learning: a review. Computer Assisted Language Learning 30, 64–90. doi:10.1080/09588221.2016.1242502

Spector, J.M., 2013. Emerging educational technologies and research directions. Journal of Educational Technology & Society 16, 21.

Steenbergen-Hu, S., Cooper, H., 2013. A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. Journal of Educational Psychology, Advanced Learning Technologies 105, 970–987. doi:10.1037/a0032447

Toetenel, L., Rienties, B., 2016. Analysing 157 learning designs using learning analytic approaches as a means to evaluate the impact of pedagogical decision making. British Journal of Educational Technology 47, 981–992. doi:10.1111/bjet.12423

Vandewaetere, M., Clarebout, G., 2014. Advanced Technologies for Personalized Learning, Instruction, and Performance, in: Spector, J.M., Merrill, M.D., Elen, J., Bishop, M.J. (Eds.), Handbook of Research on Educational Communications and Technology. Springer New York, pp. 425–437. doi:10.1007/978-1-4614-3185-5_34

Vie, J.-J., Popineau, F., Bruillard, É., Bourda, Y., 2017. A Review of Recent Advances in Adaptive Assessment, in: Peña-Ayala, A. (Ed.), Learning Analytics: Fundaments, Applications, and Trends, Studies in Systems, Decision and Control. Springer International Publishing, pp. 113–142. doi:10.1007/978-3-319-52977-6_4

Wang, S., He, F., Andersen, E., 2017. A Unified Framework for Knowledge Assessment and Progression Analysis and Design, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17. ACM, New York, NY, USA, pp. 937–948. doi:10.1145/3025453.3025841

Wilson, K., Zack, N., 2015. The Knewton Platform. A General-Purpose Adaptive Learning Infrastructure, Knewton.

Yahya, A.A., Osman, A., Khairi, M., 2016. Mining Educational Data to Analyze Teaching Effectiveness. Journal of Theoretical and Applied Information Technology 89, 267.