What can innovation in engineering education do for you as a student and what can you do as a student for innovation in engineering education?

Carlos Alario-Hoyos
Universidad Carlos III de Madrid
calario@it.uc3m.es

Abstract

Innovation in education in general, and innovation in engineering education in particular, must be supported by data, properly collected and analyzed to guide decision-making processes. Today it is possible to collect data from many more stakeholders (not just students), and also to collect much more data from each stakeholder. Nevertheless, low-level data collected by monitoring the interactions of the multiple stakeholders with learning platforms and other computing systems must be transformed into meaningful high-level indicators and visualizations that guide decision-making processes. The aim of this paper is to discuss some notable trends in data-driven innovation in engineering education, including: 1) improvement of educational content; 2) improvement of learners' social interactions; 3) improvement of learners' self-regulated learning skills; and 4) prediction of learners' behavior. However, there are also significant risks associated with data collection and processing, including privacy, transparency, biases, misinterpretations, etc., which must also be taken into account, and that require creating specialized units, and training the personnel in data management.

Keywords: *Data-driven innovation, learning analytics, digital education.*

1. Introduction

The rapid changes taking place in today's world have important consequences for education in general, and for engineering education in particular. Traditional "chalk and talk" teaching methodologies are not effective in engineering education, where it is necessary to combine the explanation of complex concepts with the practice of these concepts in realistic scenarios. In order to do this, it is necessary to apply learner-centered methodologies that promote active learning (Alario-Hoyos, Estévez-Ayres, Delgado Kloos, Villena-Román, Muñoz-Merino, et al., 2019), and that can be tailored to different education contexts, including face-to-face education in the classroom, online education, and blended (or hybrid) education (Pérez-Sanagustín, Hilliger, Alario-Hoyos, Delgado Kloos, & Rayyan, 2017).

Innovation in engineering education cannot take place without the support of a multidisciplinary team of specialized professionals. Instructional designers must advise teachers on how to redesign their traditional course to turn it into a blended learning experience, making the most of the available resources with the aim to promote active learning (Baepler, Walker, & Driessen, 2014). Pedagogues and psychologists must contribute with the understanding and development of self-regulated learning (SRL) skills to ensure student success, especially in engineering education where further development of SRL skills is required to succeed (Zimmerman, 2013). Developers must create applications and simulators to put into practice and evaluate the key concepts of each course. Researchers and data scientists must collect and analyze data, offering visualizations so that the main stakeholders of the educational process (students, teachers, managers, etc.) can make informed decisions. All in all, innovation in engineering education is a multidisciplinary effort in which all supporting professional and actions must be aligned with the ultimate goal to benefit students.

Among the abovementioned professionals, the role of researchers and data scientists is becoming more and more important in the last few years. Over the course of time, many efforts to innovate in engineering education have been undertaken without considering the research advances in the field of technology-enhanced learning. Today, with the amount of data that can be collected for the particular context of each educational institution, there is no longer an excuse for not implementing data-driven decision-making processes at the different levels: 1) teachers must rely on data to improve their content and the methodologies used in their classes; 2) students must rely on data to detect their knowledge gaps and improve their SRL skills; and 3) managers must rely on data to detect unbalanced programs or unsatisfactory courses, and to offer personalized pathways for students.

The aim of this paper is to present and discuss some relevant trends in data-driven innovation in engineering education on the basis of recent research results. All the trends analyzed here have the student as the ultimate target stakeholder, so these trends are framed within the area known as learning analytics (Ferguson, 2012). But the collection and processing of data for decision-making has a twofold side and is not without risks, some of which are also discussed in this paper. This is precisely the structure followed in the rest of this paper, with section 2 discussing trends on data-driven innovation in engineering education, section 3 dedicated to discussing associated risks, and section 4 presenting the conclusions of this work.

2. Some trends on data-driven innovation in engineering education

One of the most important trends in data-driven innovation for engineering education is the improvement of educational content. Educational content may include, for example, video lectures, automatic correction exercises, and other additional resources (texts, animations, simulations, etc.). Nowadays, it is possible to collect low-level data related to the interaction of each learner with each educational resource, including (Ruipérez-Valiente, Muñoz-Merino, Leony, & Delgado Kloos, 2015): when the learner starts to play a video, when the learner stops a video, number of seconds of the video watched by the learner, when the learner attempts an automatic correction exercise, number of attempts in each automatic correction exercise by the learner, learning sequence followed by the learner when moving between educational resources, etc. With this low-level data it is possible to detect videos or exercises that are not working correctly and that need to be improved. Some high-level indicators that contribute to detecting video lectures that need to be revised are: a high number of repetitions in a fragment of a video (which typically denotes a complex explanation or an error in that fragment), and a high percentage of students who do not watch a video from the beginning to the end (which typically denotes inappropriate content for the student's level or lack of engagement in the teacher's explanation). Some high-level indicators that contribute to detecting automatic correction exercises that need to be revised are: a very low number of incorrect answers in formative exercises (which typically denotes very simple exercises that may cause boredom and waste students' time); and a very high number of incorrect answers in summative exercises (which typically denotes very complex exercises that may cause frustration as students are not well prepared to solve those exercises).

Another important trend in data-driven innovation in engineering education refers to the improvement of social interactions among learners, and applies typically to courses with a very large number of students, either online or blended courses, and where teachers cannot provide personalized assistance due to the very large number of social interactions that take place in the course. Research in this line focused on characterizing the social interactions produced in a course and on proposing methods and visualizations to help teachers make decisions about how to improve their course design. For example, there have been research studies who detected that the most appropriate tool to manage social interaction in courses with a very large number of students is the built-in forum provided by the learning platforms (Alario-Hoyos, Pérez-Sanagustín, Delgado-Kloos, Parada G., & Muñoz-Organero, 2014). Some other research studies focused on the identification of leaders within the community of learners, characterizing these leaders as the most active students in the course forum (Alario-Hoyos, Muñoz-Merino, Pérez-Sanagustín, Delgado Kloos, & Parada G., 2016); this identification of leaders is important to facilitate teachers' work, as leaders can act as a bridge between faculty and the rest of the students, even receiving special roles to be able to curate forum messages. Some other research studies focused on analyzing the overall class mood from social interactions, calculating the polarity of messages (positive, neutral, negative) posted by students in the course forum; the polarity of messages was calculated by applying word dictionaries and syntax rules, and the aim was to detect parts of the course in which the overall class mood was more positive or more negative to take corrective measures in the second case (Moreno-Marcos, Alario-Hoyos, Muñoz-Merino, Estévez-Ayres, & Delgado Kloos, 2018). Finally, all the data collected from social interactions can be used as input to develop chatbots or conversational agents programmed to give support to students in specific courses (Delgado Kloos, Catalán-Aguirre, Muñoz-Merino, Alario-Hoyos, 2018).

An additional relevant trend in data-driven innovation in engineering education is the study, characterization, and support of students' development of self-regulated learning (SRL) skills. SRL skills are particularly important in engineering education due to, among other things, the complexity of the contents and the permanent retraining demanded in today's engineers. Appropriate strategies to self-regulate each one's learning should be applied in every learning stage (before, during and after each learning activity), including, for instance, setting reasonable and measurable objectives (before), seeking help when necessary (during), self-reflecting on the work done and the objectives achieved (after), etc. (Alonso-Mencía, Alario-Hoyos, Maldonado-Mahauad, Estévez-Ayres, Pérez-Sanagustín, & Delgado Kloos, 2019). The study and characterization of SRL skills in engineering courses led to the conclusion that strategies related to appropriate time management are the most problematic ones for learners. Therefore, specific interventions need to be done to facilitate time management, both when designing a course and when developing support tools (Alario-Hoyos, Estévez-Ayres, Pérez-Sanagustín, Delgado Kloos, & Fernández-Panadero, 2017). It is also important to support students' self-reflection by offering high-level visualizations based on low-level data with the aim to increase students' awareness on the desired level to be achieved and that of their classmates, both for an entire course and for each module or part of a course (Ruipérez-Valiente, Muñoz-Merino, Leony, & Delgado Kloos, 2015).

One last, but not least, relevant trend in data-driven innovation in engineering education is the prediction of students' behavior from previously collected data and appropriately trained machine learning models. The variables that are typically predicted through these models refer to students' partial or final grades in a course, and also to whether a student will abandon a course on not, and, by extension, whether a student will abandon a complete study program (bachelor's degree or master's degree) or not. The aim of these prediction models is to take corrective measures in order to prevent students from failing a course or from dropping out of a course or study program. Studies on prediction in education have detected that, in general, low-level data, such as the interaction with educational content (e.g., videos and exercises) have greater predictive power than data collected from self-reported questionnaires (such as students' intentions and motivations) (Moreno-Marcos, Alario-Hoyos, Muñoz-Merino, & Delgado Kloos, 2018). There are still important gaps in this research line, including: proposing generalizable models applicable to different educational contexts and areas of knowledge; and developing predictive models and tools for real-time data collection and processing in order to improve the implementation of corrective measures.

3. Some risks of data-driven innovation in engineering education

The collection and processing of data for decision making in engineering education is not without problems. In fact, there are important risks to worry about, which have been highlighted by numerous experts on learning analytics, educational data mining and data protection policies, (Dringus, 2012) (Khalil, Taraghi, & Ebner, 2016), as well as certain ethical considerations (Slade, & Tait, 2019). Some of these risks are briefly described in this section, although each of them would deserve an entire paper for discussion.

A first problem refers to the secure storage of collected data. Many institutional education systems are not prepared to store large amounts of data in a secure way and are likely to have breaches that can compromise important data, including students' personal data. Relying on third-party services for data storage in the cloud can also lead to an inappropriate use of the data collected by these service providers.

A second problem refers to data privacy. It is very important to have an institutional strategy that clearly states what data needs to be collected from students and who has access to the data. Many educational institutions are not aware of all the data they collect (or can potentially collect), do not have mechanisms to control who has access to the data collected, and/or do not have an internal policy to facilitate that the right people can make proper informed decisions being able to access the data they are entitled to.

A third problem refers to always getting explicit consent from the students for collecting data. It is very important that students (who actually own the data) know at all times what data are being collected from them and for what purpose the data are going to be used. In addition, there are international laws such as the GDPR (General Data Protection Regulation) in Europe that require, among other things, the removal of data collected upon request by its owner. Many educational institutions do not have data collection ethics committees and are not prepared to delete collected data upon request.

A fourth problem refers to transparency, or rather lack of transparency, of many algorithms and systems which are private and whose code is not open. The lack of transparency prevents algorithms and systems from being audited to better understand how they work. Even if one relies on third-party algorithms and systems to make informed decisions it is important to know how the results and visualizations they provide were obtained.

A fifth problem refers to bias. Many artificial intelligence systems use data collected in the past to make their calculations and predictions. However, data collected in the past may have important biases, such as a gender imbalance, which could lead to promoting more male students versus female students or vice versa. In general, it is important to take minorities into account when using data from the past as input so that these are not penalized.

Finally, it is important to bear in mind that humans may misinterpret the results and visualizations obtained from the data processed. Sometimes there are people who deliberately twist the data to fit a pre-designed theory, instead of discarding this theory if data advise to do so. Actually, from a very large dataset, and taking only a subset of it, erroneous and unreproducible conclusions can be very easily reached.

4. Conclusions

Innovation in engineering education must be informed by data. Teachers must be aware of their students' performance (individually and at the class level) to support those who need more help as well as offer top quality educational resources progressively improved according to the data collected. Students should define a curriculum adapted to their particular needs and develop their SRL skills. Institutions must create specialized units for data collection and processing and adequately train their staff (including teachers) for a data management culture. Nevertheless, there are numerous risks to be aware of, and a trade-off is needed so that these risks do not slow down innovation in educational institutions.

Acknowledgements

The author acknowledges the support from Erasmus+ Capacity Building in the Field of Higher Education, project InnovaT (598758-EPP-1-2018-1-AT-EPPKA2-CBHE-JP).

References

Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Parada G., H. A., & Muñoz-Organero, M. (2014). Delving into participants' profiles and use of social tools in MOOCs. *IEEE Transactions on Learning Technologies, 7*(3), 260-266.

Alario-Hoyos, C., Muñoz-Merino, P. J., Pérez-Sanagustín, M., Delgado Kloos, C., & Parada G., H. A. (2016). Who are the top contributors in a MOOC? Relating participants' performance and contributions. *Journal of Computer Assisted Learning, 32*(3), 232-243.

Alario-Hoyos, C., Estévez-Ayres, I., Pérez-Sanagustín, M., Delgado Kloos, C., & Fernández-Panadero, C. (2017). Understanding learners' motivation and learning strategies in MOOCs. *Intern. Review of Research in Open and Distributed Learning, 18*(3), 119-137.

Alario-Hoyos, C., Estévez-Ayres, I., Delgado Kloos, C., Villena-Román, J., Muñoz-Merino, P. J., & Llorente-Pérez, E. (2019). Redesigning a freshman engineering course to promote active learning by flipping the classroom through the reuse of MOOCs. *International Journal of Engineering Education, 35*(1), 385-396.

Alonso-Mencía, M. E., Alario-Hoyos, C., Maldonado-Mahauad, J., Estévez-Ayres, I., Pérez-Sanagustín, M., & Delgado Kloos, C. (2019). Self-regulated learning in MOOCs: lessons learned from a literature review. *Educational Review (published online)*, 1-27.

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. Computers & Education, 78, 227-236.

Delgado Kloos, C., Catalán-Aguirre, C., Muñoz-Merino, P. J., Alario-Hoyos, C. (2018). Design of a Conversational Agent as an Educational Tool, In *Learning with MOOCs 2018 (LWMOOCS V*) (pp. 27-30). IEEE

Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks, 16*(3), 87-100.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5/6), 304-317.

Khalil, M., Taraghi, B., & Ebner, M. (2016). Engaging Learning Analytics in MOOCs: the good, the bad, and the ugly. In *Proceedings of the International Conference on Education and New Developments (END)* (pp. 3-7).

Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., & Delgado Kloos, C. (2018). Prediction in MOOCs: A review and future research directions. *IEEE Transactions on Learning Technologies (published online)*. 1-19.

Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estévez-Ayres, I., & Delgado Kloos, C. (2018). Sentiment Analysis in MOOCs: A case study. *In 2018 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1489-1496). IEEE.

Pérez-Sanagustín, M., Hilliger, I., Alario-Hoyos, C., Delgado Kloos, C., & Rayyan, S. (2017). H-MOOC framework: reusing MOOCs for hybrid education. *Journal of Computing in Higher Education, 29*(1), 47-64.

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Delgado Kloos, C. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior, 47,* 139-148.

Slade, S., & Tait, A. (2019). *Global guidelines: Ethics in learning analytics*. Oslo: International Council for Open and Distance Education (ICDE).

Zimmerman, B. J. (2013). Theories of self-regulated learning and academic achievement: An overview and analysis. In *Self-regulated learning and academic achievement* (pp. 10-45). Routledge.