Auditorium
1.35-2.20pm

NUS's vision of learning analytics, individual autonomy and data privacy

Robert KAMEI[1]*, Kiruthika D. RAGUPATHI[2], Alice CHRISTUDASON[3] and Kevin HARTMAN[1]
[1] Institute for Applied Learning Science and Educational Technology, National University of Singapore
[2] Centre for the Development of Teaching and Learning, National University of Singapore
[3] School of Design and Environment, National University of Singapore
*pvorkk@nus.edu.sg

Session Objective

In this panel discussion, Kiruthika Devi Ragupathi, Alice Christudason and Robert Kamei will provide their views on learning analytics, individual autonomy, and data privacy as they relate to the National University of Singapore's (NUS) position as a public good and a world class institution. The session will begin and end with an online poll of the audience. After the initial poll, the panellists will link the session's themes together in a series of short talks. Questions will then be taken directly from the audience.

Background

If the past generation of educational researchers sought to determine "what works" (National Research Council, 2000), the current generation of educational researchers is seeking to determine "what works for whom." By harnessing big data practices, machine learning, and predictive analysis, learning analytics holds the promise of reliability diagnosing learning needs and matching those needs to a best course of treatment. Learning analytics is defined as the collection, processing, visualisation, and interpretation of data gleaned from past learning activities for use in making decisions about future learning activities (Siemens & Long, 2011). The discipline spans measuring learning during standalone activities (Cutumisu, Blair, Chin, & Schwartz, 2015) to evaluating the effectiveness of teaching and learning of entire institutions (Macfadyen, Dawson, Prest, & Gašević, 2016). Two key questions woven throughout learning analytics research are 1) How much autonomy should learners be given when making decisions about their learning, and 2) How personal should the data used to make those decisions be?

Learning Analytics and Individual Autonomy

In university settings, learning analytics has already been used to identify students at risk of stopping out (Taylor, Veeramachaneni, & O'Reilly, 2014)—suddenly ceasing course activities—and dropping out (Colvin *et al.*, 2015)—withdrawing from the university. In these contexts, systems have been developed to automatically contact at-risk students with reminders about assignments and promises of support when they show certain indicators. From an autonomy perspective, these interventions appear fairly innocuous. When a student misses a deadline, the system generates a message to the student. How that student responds to the message is still within his or her control. However, what happens if the system starts using risk-factors to make decisions about how materials are packaged, which version of an activity is delivered and how assessments are evaluated without informing learners? Such a system could already be developed with the data NUS administrators and researchers already make available in the NUS Data Lake. The NUS Data Lake contains anonymised records of NUS students, their demographic details, their past course performance, their current course enrolment, and every click they have committed on IVLE. Couple that information with additional measures collected from surveys on motivation, self-directed learning and personality and suddenly the university can choose to micro-target students on any number of dimensions even without knowing their names. One of the questions for the panel's discussion is given what NUS could do, how should we determine where the limits should be?

Learning Analytics and Data Privacy

As universities learn more about their students and build better profiles of their learners, they put their students and instructors at greater risk of data exposure and re-identification. Data exposure occurs when sensitive internal data makes its way into a more public sphere (Ponemon Institute, 2013). Data exposure can occur through hacking but it occurs more often through the interaction between the university and third-party vendors or even researchers. The more information that is catalogued and linked together, the more valuable that information becomes to researchers, instructors, students, and people outside the NUS community. The greater the number of people working with a data set, the greater the chance that data may become exposed. Even though the data in the NUS Data Lake is anonymised, we have learned that despite this protection, someone who has access to enough data could re-identify individuals. EdX researchers identified the risk of letting data users systematically query databases in ways that allow them to re-identify students and research participants (Daries *et al*., 2014). Their recommendation was to pay more attention to the design of analytics systems to prevent and track re-identification. The easier we make it for researchers to use NUS data, the easier we make it for someone to use this data for intended purposes. What is the right balance for NUS?

Conclusion

Individual autonomy and data privacy are two key issues NUS is discussing with regard to learning analytics and the future of learning. As the university knows more about the needs and interests of its learners, it has the power to transform learning experiences, to cultivate inquiring minds, and reducing the barriers to taking research-informed approaches. At the same time, the university needs to better define where the boundaries are for what can and cannot be ethically done. We hope to facilitate the continuation of this ongoing conversation with the proposed discussion panel.

Keywords

Scholarship of teaching and learning, educational technology, learning analytics

References

Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S., & Fisher, J. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Sydney: Australian Office for Learning and Teaching.

Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2015). Posterlet: A game-based assessment of children's choices to seek feedback and to revise. *Journal of Learning Analytics*, *2*(1), 49-71. https://doi.org/10.18608/jla.2015.21.4

Daries, J. P., Reich, J., Waldo, J., Young, E. M., Whittinghill, J., Ho, A. D., ... & Chuang, I. (2014). Privacy, anonymity, and big data in the social sciences. *Communications of the ACM*, *57*(9), 56-63. http://dx.doi.org/10.1145/2643132

Macfadyen, L. P., Dawson, S., Prest, S., & Gašević, D. (2016). Whose feedback? A multilevel analysis of student completion of end-of-term teaching evaluations. *Assessment & Evaluation in Higher Education*, *41*(6), 821-839. http://dx.doi.org/10.1080/02602938.2015.1044421

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academies Press.

Ponemon Institute. (2013, May). *2013 Cost of Data Breach Study: Global analysis*. Retrieved from: http://www.symantec.com/content/en/us/about/media/pdfs/b-cost-of-a-data-breach-global-report-2013.en-us.pdf.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 30. Retrieved from https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education.

Taylor, C., Veeramachaneni, K., & O'Reilly, U. M. (2014). Likely to stop? Predicting stopout in massive open online courses. Retrieved from https://arxiv.org/abs/1408.3382v1.