

Towards Flexible MOOCs: Student-Sourcing of Learning Content at Scale

MOOCs (massive open online courses) have emerged as an important component of modern online learning. Among the most exciting opportunities brought by MOOC is the ability to leverage the power of Big Data to improve learning. Indeed, for the first time MOOCs offered a chance to log data of thousands of learners taking the same course on a relatively fine-grain level. The availability of this data have been discussed in both public media and research literature producing a numbers of promises on how this data can be used and how it could change learning. Now, after more than 5 years of research we can see that a numbers of these promises were fulfilled. We can now understand the learning process better as well as establish more reliable connections between students' characteristics, course performance, and learning outcomes. Yet, at least in one aspect we are not much closer to the goal than we were 5 years ago.

One of the promises associated with MOOC-level learning data was *data-driven personalized sequencing* of learning content - the ability to suggest most beneficial learning activity for every student at every moment of learning. While personalized sequencing is not a new idea, traditional sequencing approaches are expensive to implement since they are based on expert knowledge about learning content and process. MOOC-level data opened a chance to learn personalized sequencing rules directly from data. Consider a simple example: that data shows that a category of students experiences serious problems when solving a specific problem unless they had a chance to explore an example shortly before attempting this problem. In a personalized version of the course, the students in this category could be recommended to work with this example before the problem. While many data-driven sequencing ideas have been shared, the attempts to implement it faced a critical obstacle: the lack of diverse learning content. To fulfill its promise data-driven sequencing needs a considerable variety of learning content so that every learner can receive the best item to work with. As it appears, modern online learning systems have a very thin layer of content, basically, just one path that everyone has to take. In this context, attempts to learn "the best" path from student data result in re-learning the same single road.

To fulfill the data-driven personalization promise we need a large variety of learning content – but where it can come from? It is a considerable effort for MOOC authors to create even the single path. The paper "Adaptive Social Learning Based on Crowdsourcing" by Evgeny Karataev and Vladimir Zadorozhny (IEEE Transactions on Learning Technologies, Volume: 10, Issue: 2, April-June 1 2017) suggests to leverage the strongest side of online learning systems – their learners. Crowdsourcing (or more exactly *learner-sourcing*) of learning content has been explored in the past, however past work on learner-sourcing have not been motivated by the need to produce a larger variety of content for personalization. Karataev and Zadorozhny bring the ideas of learner-sourcing and personalization together by introducing crowdsourcing on the level of self-contained *lesslets* and suggesting a novel framework that supports both, organizing lesslets and learning

efficient pathways through the lesslet universe from student data. Their paper presents details of their social learning framework and reports the result of its first evaluation with data collected over several years of its use. The evaluation specifically explored the emergence of “personal pathways” and the performance of some content recommendation algorithms.

While this work could be considered as an early attempt to connect learner-sourcing and data-driven personalization, we believe that it indicates some important direction of future research. We hope that more work in this direction will follow breaking the “content diversity” obstacle of modern learning environments.

[Peter Brusilovsky](#)

University of Pittsburgh

Peter Brusilovsky is a professor of Information Science and Intelligent Systems at the University of Pittsburgh, Contact him at peterb@pitt.edu.