

Learnersourcing: Student-generated Content @ Scale

Steven Moore
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
stevenjamesmoore@gmail.com

John Stamper
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
john@stamper.org

Christopher Brooks
University of Michigan
Ann Arbor, Michigan, USA
broosch@umich.edu

Paul Denny
The University of Auckland
Auckland, New Zealand
paul@cs.auckland.ac.nz

Hassan Khosravi
The University of Queensland
Brisbane, QLD, Australia
h.khosravi@uq.edu.au

1 Background

Involving students in the process of creating questions, hints, examples, and other instructional content has been shown to benefit student learning [1]. This is known as a form of *learnersourcing*,

where students complete activities in online courses that produce content that can be leveraged by future learners [2]. Learnersourcing offers a domain agnostic way to help scale the creation of high-quality assessments, while also helping students learn the course content. Involving students in this process, instead of crowdworkers or other accessible parties, is essential due to the domain knowledge students are often required to possess by being enrolled in a particular course. It also offers the unique perspective of someone more novice that might have unique insights due to actively learning the material, overcoming potential expert blind spot that instructors typically have when creating assessments [3].

Several learnersourcing systems have been developed to support student participation in a variety of activities, including content review and problem explanation [4]. Two popular learnersourcing systems, PeerWise [5] and RiPPLE [6], were developed to assist students in the process of generating and reviewing multiple-choice questions. Using these two tools, tens of thousands of students have generated and reviewed millions of multiple-choice questions. This includes the students' providing assessments of the question quality and feedback on how to potentially improve them for the authors and others to utilize. These tools are domain agnostic and can be incorporated into any existing course, allowing for ample research and educational opportunities. Previous research has shown that the use of these tools enables students to generate high-quality assessments that can be leveraged by the students as practice opportunities, all while being effective for their learning [7]. Learnersourcing activities can also be directly implemented into existing courseware, such as a MOOC platforms like Coursera

*Article Title Footnote needs to be captured as Title Note

†Author Footnote to be captured as Author Note

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WOODSTOCK '18, June, 2018, El Paso, Texas USA

© 2018 Copyright held by the owner/author(s). 978-1-4503-0000-0/18/06...\$15.00

or edX, or learning management systems like Canvas or Blackboard, without the use of external tools. Recent learnersourcing research has demonstrated that students can effectively generate quality questions, even when the learnersourcing activities have limited scaffolding and student participation with them is optional [8,9].

The past five years of L@S has included papers on learnersourcing, as the field continues to grow in popularity, and new techniques are developed to improve the quality of student-generated content. We want to share this concept of learnersourcing with the broader audience of the conference, to show them how they can leverage these existing tools or create activities within their own courses and research platforms to enable students to generate such content. These systems and learnersourcing activities, along with several datasets collected from their use by thousands of students, are available for instructors and researchers to use. Learnersourcing provides invaluable data that can be used to generate question banks, help answer questions related to student learning, and provides ample opportunities for machine learning and natural language processing work.

The main goal of this workshop is to expose attendees to the ample opportunities in the learnersourcing space, including instructors, researchers, learning engineers, and many other roles. We believe participants from a wide range of backgrounds and prior knowledge on learnersourcing can both benefit and contribute to this workshop. As the learnersourcing process involves many stakeholders (students, instructors, researchers, instructional designers, etc.), multiple viewpoints can help to inform what future student-generated content might be useful, new and better ways to assess the quality of the content, and spark potential collaboration efforts between attendees. We ultimately want to show how everyone can make use of learnersourcing and have participants gain hands on experience using these tools, creating their own learnersourcing activities using them or their own platforms, and take part in an exploration of the learnersourcing data we share out. Our hope is to attract attendees interested in scaling the generation

of instructional and assessment content and those interested in the use of online learning platforms.

2 Pre-Workshop Plans

Prior to the workshop, we will provide access to the two core systems being discussed at the beginning of the session. Participants will be able to investigate the activities and affordances each system offers, the review process for questions, and much more to both learn about and even participate in learnersourcing themselves. Additionally, we will make several datasets collected from these systems available to the participants, that they are free to explore and utilize during the workshop or for future endeavors. Our hope is that during the workshop, they can then ask any questions and gain a better understanding of the types of learnersourcing data, if they wish to do so, before we meet. Finally, we will post a brief survey to collect the backgrounds and interests of the participants to help tailor our discussions and activities.

3 Workshop Structure

This will be a half-day workshop. The workshop focus will be on examining the tools, processes, and content that is both used and generated through learnersourcing. We will begin with introductions and an overview of the learnersourcing landscape, to bring all participants, regardless of background, up to speed on the concept and latest trends. Three presentations and demos will then be ran to highlight different learnersourcing tools, with an emphasis on how the student-generated content can be used by instructors and researchers. Participants will then engage in a discussion with the organizers regarding the types of student-generated content that might be the most useful in their work and how we might better engage students and assess their learning as they participate in these learnersourcing activities. From there, we will shift the discussion to the challenges and opportunities of assessing the quality of the student-generated content, which is a main research thrust in the learnersourcing community. Following this we will work with participants hands on to investigate several learnersourcing datasets collected from the aforementioned tools and explore the different questions that can be answered using this data. The workshop will conclude with a summary of the day's events, core challenges and opportunities we addressed in the discussions, and a strong emphasis on future collaborations.

We propose the following schedule:

08:30 - Introductions
 09:00 – PeerWise demo
 09:30 – RiPPLE demo
 10:00 – Coffee break
 10:15 – Discussion: Activity types and generation
 10:45 – Hands on: Adding learnersourcing activities of your own
 11:15 – Assessing learnersourcing: challenges and opportunities
 11:45 – Data sharing: Leveraging data from all three tools
 12:15 – Closing remarks and the future of learnersourcing

4 Post-Workshop Plans

We plan for the interactions during the workshop to result in the adoption of learnersourcing for many of the participants, whether that be using one of the tools, the provided datasets, or creating learnersourcing activities in their own platforms and courses. Upon completion of the workshop, we will offer participants the chance to join a Slack channel and mailing list dedicated to sharing out advances in learnersourcing. Through these channels, we will continue to share datasets collected from these and other learnersourcing systems. We envision that these datasets can be leveraged by participants for future studies and potentially be the focus of a future workshop or competition at L@S. Ultimately, we want to keep the participants involved and promote collaboration between attendees. We expect this workshop to be repeated and become part of the basis for a community of researchers who are interested in learnersourcing and assessment generation at scale.

5 Call for Participation

Involving students in the process of creating questions, hints, examples, and other instructional content has been shown to benefit student learning. This is known as a form of *learnersourcing*,

where students complete activities in online courses that produce content that can be leveraged by future learners. Learnersourcing offers a domain agnostic way to help scale the creation of high-quality assessments. There are several tools available for instructors, researchers, and practitioners alike to involve students in the process of generating assessments, such as multiple-choice questions. Through the use of these tools and learnersourcing activities, tens of thousands of students have generated millions of practice questions across a variety of domains.

In this workshop, we will demo the use of two learnersourcing tools that can be readily utilized in any course. We will also work hands on with participants to create activities in their own systems/courses and begin exploring several learnersourcing datasets made available to attendees. This workshop will bring together participants to discuss new learnersourcing activity types, methods for assessing the quality of student-generated content, and ways we might better scale learnersourcing efforts to create usable instructional and assessment materials. Participants of all backgrounds are welcome, as learnersourcing involves many stakeholders including students, instructors, researchers, practitioners, and many more!

6 Organizers

The workshop will be organized by a team of learnersourcing researchers and practitioners, all who are involved with the development of a learnersourcing systems or data repositories. Additionally, they all have prior experience conducting learnersourcing research and past workshop organization.

Steven Moore is a PhD student in Human-Computer Interaction at Carnegie Mellon University and is advised by Dr. John Stamper. His research is focused on engaging students in the learnersourcing process and finding ways to assess the quality of their contributions.

John Stamper is an Associate Professor at the Human-Computer Interaction Institute at Carnegie Mellon University and the Technical Director of the Pittsburgh Science of Learning Center DataShop. His work involves leveraging educational data mining techniques and the creation of data tools that can be used with learnersourcing data.

Christopher Brooks is an Assistant Professor at the University of Michigan and is an applied Computer Scientist who builds and studies the effects of educational technologies in higher education and informal learning environments. He has led learnersourcing efforts on the Coursera platform, where he investigated student choice in the generation of multiple-choice questions.

Paul Denny is an Associate Professor in Computer Science at the University of Auckland, New Zealand. He leads the PeerWise project, which hosts more than six million practice questions, with associated solutions and explanations, created by students from 90 countries.

Hassan Khosravi is a Senior Lecturer in the Institute for Teaching and Learning Innovation and an Affiliate Academic in the School of Information Technology and Electrical Engineering at the University of Queensland. He has conducted extensive learnersourcing research and leads the development and dissemination efforts of the RiPPLE system.

KEYWORDS

Learnersourcing, student-generated content, question creation, assessment, multiple-choice question, student learning

REFERENCES

- [1] Khosravi, H., Demartini, G., Sadiq, S., & Gasevic, D. (2021, April). Charting the design and analytics agenda of learnersourcing systems. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 32-42).
- [2] Kim, J. (2015, May). Learnersourcing: improving learning with collective learner activity. PhD Dissertation. Massachusetts Institute of Technology
- [3] Nathan, M. J., Koedinger, K. R., & Alibali, M. W. (2001, April). Expert blind spot: When content knowledge eclipses pedagogical content knowledge. In *Proceedings of the third international conference on cognitive science* (Vol. 644648).
- [4] Williams, J. J., Kim, J., Rafferty, A., Maldonado, S., Gajos, K. Z., Lasecki, W. S., & Heffernan, N. (2016, April). Axis: Generating explanations at scale with learnersourcing and machine learning. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale* (pp. 379-388).
- [5] Denny, P., Hamer, J., Luxton-Reilly, A., & Purchase, H. (2008, September). PeerWise: students sharing their multiple-choice questions. In *Proceedings of the fourth international workshop on computing education research* (pp. 51-58).
- [6] Khosravi, H., Kitto, K., & Williams, J. J. (2019). RiPPLE: A Crowdsourced Adaptive Platform for Recommendation of Learning Activities. *Journal of Learning Analytics*, 6(3), 91-105.
- [7] Duret, D., Christley, R., Denny, P., & Senior, A. (2018). Collaborative learning with PeerWise. *Research in Learning Technology*, 26, 1-13.
- [8] Moore, S., Nguyen, H. A., & Stamper, J. (2021, June). Examining the Effects of Student Participation and Performance on the Quality of Learnersourcing

Multiple-Choice Questions. In *Proceedings of the Eighth ACM Conference on Learning@ Scale* (pp. 209-220).

- [9] Singh, A., Brooks, C., Lin, Y., & Li, W. (2021, June). What's In It for the Learners? Evidence from a Randomized Field Experiment on Learnersourcing Questions in a MOOC. In *Proceedings of the Eighth ACM Conference on Learning@ Scale* (pp. 221-233).