# Partnerships for Cocreating Educational Content

**Steven Moore[1], Anjali Singh[2], Hassan Khosravi[3], Paul Denny[4], Chris Brooks[2], John Stamper[1]**
Carnegie Mellon University[1], University of Michigan[2], University of Queensland[3], University of Auckland[4]
StevenJamesMoore@gmail.com

**ABSTRACT**: We propose the first annual workshop on Partnerships for Cocreating Educational Content. This in-person workshop explores ample opportunities in leveraging humans, AI, and learning analytics to generate content, particularly appealing to instructors, researchers, learning engineers, and many other roles. The process of humans and AI cocreating educational content involves many stakeholders (students, instructors, researchers, instructional designers, etc.), thus multiple viewpoints can help to inform what future generated content might be useful, new and better ways to assess the quality of such content and to spark potential collaborative efforts between attendees. We ultimately want to show how everyone can leverage recent advancements in learnersourcing, AI, and learning analytics, and engage all participants in shaping the landscape of challenges and opportunities in this space. Our hope is to attract attendees interested in scaling the generation of instructional and assessment content and those interested in online learning platforms.

**Keywords**: Educational content creation, human-AI partnerships, learnersourcing

## 1 BACKGROUND

Globally, as educational delivery continues to transition towards online platforms hastened by the COVID-19 pandemic, the need for scalable and effective assessments has emerged as a pressing issue for instructors and educators. Amid many other logistical issues that arise from emergency online education (Hodges et al., 2020), instructors often find themselves having to generate large banks of resources such as practice and assessment questions to accommodate this new learning format. The continual creation and improvement of assessment items allows for a greater breadth of topic coverage, helps to identify well-constructed and valid assessments, and as a result, enables improved learning analytics. However, instructors and teaching staff rarely have the time or incentive to develop quality questions for formative assessments that are often used for personalization and adaptive learning; instead their efforts are often focused on creating high-stakes assessments such as quiz or exam questions (Jones, 2019). This challenge motivates the need for supporting the efforts of educational content creation via partnerships that involve pairings of instructors, students, and AI.

Partnerships for cocreating educational content often involve four distinct and iterative phases: creation, evaluation, utilization, and instructor/expert oversight. A popular student-student and student-instructor partnership that is widely becoming adopted and involves all of these phases is learnersourcing. Learnersourcing involves students generating their own educational resources and content that can be leveraged by future learners (Khosravi et al., 2021). This offers a domain agnostic way to help scale the creation of high-quality assessments, while also helping students learn the course content. Learning analytics plays a key role in this process by providing insight into how we might effectively leverage students to create educational content. This can include optimally selecting

students to create questions that target topics where those students have demonstrated expertise, as well as recommending student-generated content on specific topics to learners who are struggling with those topics.

Partnerships between student-AI and instructor-AI also present ample opportunity in terms of content creation and evaluation (Singh et al., 2022). Advances in natural language processing and generative models provide space for AI to play a fundamental role in the co-creation of content with humans or to assist with the automated evaluation of its quality. The quality evaluation of this content can be further supported by learning analytics related to how students perform on these human-AI cocreated questions, compared to traditional assessments. For instance, one newly emerging area where human oversight may be needed in the educational content creation space is with the use of generative language models. (Sarsa et al., 2022) propose the idea of robosourcing, where content generated by large language models can be used as a starting point for students to accelerate the educational content creation process. Related work has also leveraged natural language processing (Moore et al., 2020), trust-based networks (Darvishi et al., 2021), and deep learning methods (Ni et al., 2022) to assist students in the evaluation of both student- and AI-generated content. While human input remains critical in this creation and evaluation process, more work needs to look at using artificial intelligence to further support students and instructors as they create educational content.

## 2 CALL FOR SUBMISSION

While no submission is required to participate in the workshop, we encourage 2 page submissions of work-in-progress or position papers that are related to partnerships for co-creating educational content. Some related challenges are highlighted in Figure 1. When it comes to the evaluation process of having students or AI review and revise other student-generated questions, there is a challenge regarding how we can assist students in optimally acting on the provided feedback. How to best incorporate student evaluation of the materials into the learning process, such as through learner models used to power learning analytics, remains an open problem (Abdi et al., 2020). While research indicates the learning benefits of students generating questions, oftentimes the quality of student-generated questions requires improvement. Recent work demonstrated that MCQs authored by students performed as well as those authored by academics, but further work remains to investigate how we might leverage AI to assist students in making consistently high-quality learnersourced contributions (Huang et al., 2021).

Among these challenges with humans and AI cocreating educational content lie many opportunities to explore ways of making it more accessible and beneficial to student learning. A clear opportunity regarding the creation of student-generated content is the different ways we can encourage students to make high-quality contributions, such as leveraging self regulated learning interventions (Lahza et al., 2022). While much of the existing research around students and AI creating educational content involves the creation of multiple-choice questions, there are limitless activity types that can be created and evaluated using a plethora of techniques. For instance, students could work in conjunction with a large language model, like GPT-3, to develop and refine assessment questions or explanations of learning content (Sarsa et al., 2022, Moore et al., 2022). This can help them quickly improve the content they generate, while also engaging them in critical thinking as they review the model's suggestions, such as recommended distractors. On the one hand, the increasing automation supported by such models may suggest less need for human input, but there is a need for caution. In

their review of the opportunities and risks offered by foundation models, Bommasani et al. explicitly warn against the removal of teachers from such a loop (Moore et al., 2021). Large language models are trained on broad data produced by humans, and thus are known to suffer from biases similar to humans. Using automatically generated content without human oversight for educational content generation runs the risk of perpetuating some of these biases. We see a human-in-the-loop approach, involving both students and instructors, as essential for moderating biases and improving and tailoring the performance of the underlying generative models for suitability in learnersourcing contexts.

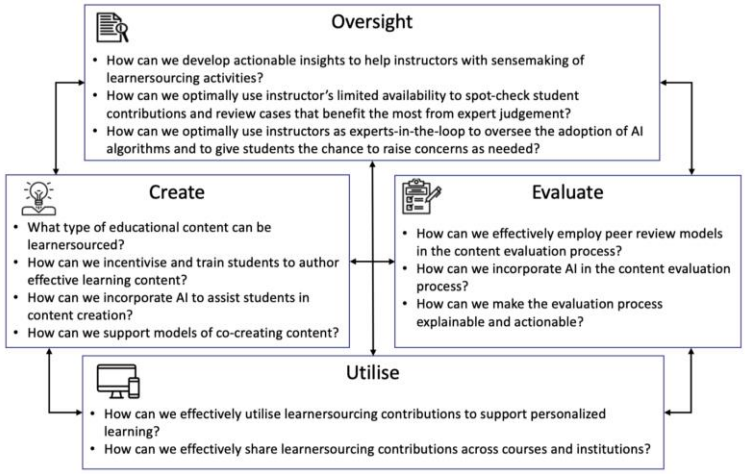


**Figure 1: Challenges and opportunities relating to the four key aspects involved in the creation of educational content involving students, instructors, and AI.**

## 3 WORKSHOP STRUCTURE

The workshop will run as an interactive half-day session with mini-presentations and round-table discussions on the theme. The provisional schedule is given below:

- **Introductions**: Introductions of workshop organizers and participants, and a background to the focus of the workshop.
- **Short Presentations**: Authors of accepted submissions present their work which would be followed by a Q&A session
- **Round-table discussion**: Participants will move around specific topics of interest related to various types of partnerships for creating educational content including partnerships between: student-student, student-AI, student-instructor, and instructor-AI.
- **Open discussion**: An open discussion will be facilitated among all participants summarizing activities from the round table discussions and building consensus using the co-creation of shared notes and resources.
- **Concluding remarks and community engagement:** Closing remarks on the workshop will be made with future steps. In addition, a Slack channel has been created to keep the participants involved and promote collaboration between attendees.

## 4 OUTCOMES

The main goal of this workshop is to explore how partnerships between students, instructors, and AI can be leveraged for creating educational content and how learning analytics informs these

interactions. We believe participants from a wide range of backgrounds and prior knowledge on learnersourcing, machine learning, and learning analytics can both benefit and contribute to this workshop. As this creation of educational content involves many stakeholders (students, instructors, researchers, instructional designers, etc.), multiple viewpoints can help to inform what future student- and AI-generated educational content might be useful, new and better ways to assess the quality of the content, and spark potential collaboration efforts between attendees. The accepted submissions will be published as part of a CEUR proceedings.

## REFERENCES

Abdi, S., Khosravi, H., & Sadiq, S. (2020). Modelling learners in crowdsourcing educational systems. *International Conference on Artificial Intelligence in Education*, 3–9.

Darvishi, A., Khosravi, H., & Sadiq, S. (2021). Employing peer review to evaluate the quality of student generated content at scale: A trust propagation approach. *Proceedings of the Eighth ACM Conference on Learning@ Scale*, 139–150.

Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). *The difference between emergency remote teaching and online learning*.

Huang, A., Hancock, D., Clemson, M., Yeo, G., Harney, D., Denny, P., & Denyer, G. (2021). Selecting student-authored questions for summative assessments. *Research in Learning Technology*.

Jones, J. A. (2019). Scaffolding self-regulated learning through student-generated quizzes. *Active Learning in Higher Education*, *20*(2), 115–126.

Khosravi, H., Demartini, G., Sadiq, S., & Gasevic, D. (2021). Charting the design and analytics agenda of learnersourcing systems. *LAK21: 11th International Learning Analytics and Knowledge Conference*, 32–42.

Lahza, H., Khosravi, H., Demartini, G., & Gasevic, D. (2022). Effects of Technological Interventions for Self-regulation: A Control Experiment in Learnersourcing. *LAK22: 12th International Learning Analytics and Knowledge Conference*, 542–548.

Moore, S., Nguyen, H. A., & Stamper, J. (2020). Evaluating Crowdsourcing and Topic Modeling in Generating Knowledge Components from Explanations. *International Conference on Artificial Intelligence in Education*, 398–410.

Moore, S., Nguyen, H. A., & Stamper, J. (2021). Examining the Effects of Student Participation and Performance on the Quality of Learnersourcing Multiple-Choice Questions. *Proceedings of the Eighth ACM Conference on Learning@ Scale*, 209–220.

Moore, S., Nguyen, H. A., Bier, N., Domadia, T., & Stamper, J. (2022). Assessing the quality of student-generated short answer questions using GPT-3. In the European conference on technology enhanced learning, 243-257.

Ni, L., Bao, Q., Li, X., Qi, Q., Denny, P., Warren, J., Witbrock, M., & Liu, J. (2022). Deepqr: Neural-based quality ratings for learnersourced multiple-choice questions. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(11), 12826–12834.

Sarsa, S., Denny, P., Hellas, A., & Leinonen, J. (2022). Automatic Generation of Programming Exercises and Code Explanations Using Large Language Models. *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1*, 27–43.

Singh, A., Brooks, C., & Doroudi, S. (2022). Learnersourcing in Theory and Practice: Synthesizing the Literature and Charting the Future. In Proceedings of the Ninth ACM Conference on Learning@ Scale (234-245).