Juho Kim • Research Statement

As a human-computer interaction (HCI) researcher, I build interactive systems powered by large-scale data from users. With an unprecedented scale of users interacting with online platforms, data generated from their interactions can be used to understand collective preferences and opinions, create new content, and improve future users' experience. In several domains, my research approach 1) extracts meaningful patterns from natural interaction traces, 2) elicits specific information from users by designing microtasks that are inherently meaningful to them, and 3) changes the interface behavior immediately as more data becomes available. These data-driven methods demonstrate how interaction data can be powerful building blocks for enhancing massive-scale learning, planning, discussion, collaboration, and sensemaking online.

Specifically, my research has focused on educational videos in learning-at-scale platforms. My primary approach has been **learnersourcing**, in which learners collectively generate novel content and interfaces for future learners while engaging in a meaningful learning experience themselves. Millions of learners today use educational videos from online platforms such as YouTube, Khan Academy, Coursera, or edX. Learners can be a qualified and motivated crowd who can help improve the content and interfaces. What if learners' collective activity could help identify points of confusion or importance in a video, reconstruct a solution structure from a tutorial, or create alternate explanations and examples? Learnersourcing can improve the content and interfaces in a way neither experts, nor computers, nor existing crowdsourcing methods can achieve at scale. **My research demonstrates that interfaces powered by learnersourcing can enhance content navigation, create a sense of learning with others, and ultimately improve learning.**

I draw on several fields to design learnersourcing applications: crowdsourcing to aggregate small contributions into meaningful artifacts; social computing to motivate participation and build a sense of community among learners; content-based video analysis techniques such as computer vision and natural language processing to complement learner input; and learning science to inform the design of learnersourcing tasks that are pedagogically meaningful. I explore two types of learnersourcing: **passive learnersourcing** uses data generated by learners' natural interaction with the learning platform, and **active learnersourcing** prompts learners to provide specific information.

Passive Learnersourcing: Natural learner interactions improve video learning

In traditional classrooms, teachers adapt their instruction to students based on their level of engagement and confusion. While online videos enable access for a wide audience, instructors and learners are disconnected; it is as if instructors are talking to a wall without feedback from learners watching the video. I created a thread of research that leverages natural learning interaction data to better understand and improve video learning, specifically using thousands of learners' second-by-second video player interaction traces (e.g., clicking the play button in the video player). I first analyzed this data post-hoc to guide the improvement of the material, and took a step further to establish a faster loop so that video presentation dynamically adapts to collective video learning patterns.

Data analysis of 39 million MOOC video clicks. Exploratory data analyses of four massive open online courses (MOOCs) on the edX platform investigated 39 million video events and 6.9 million watching sessions from over 120,000 learners. Analyzing collective in-video interaction traces revealed video interaction patterns, one of which is **interaction peaks**, a burst of play button clicks around a point in a video indicating points of interest and confusion for many learners. A key observation was that 61% of the peaks accompany visual transitions in the video, e.g., a slide view to an instructor view (Figure 1). Extending this observation, I identified student activity patterns that can explain peaks, including playing from the beginning of new material, returning to missed content, and replaying a brief segment [1]. A further analysis investigated how video production factors affect video engagement: shorter (less than six

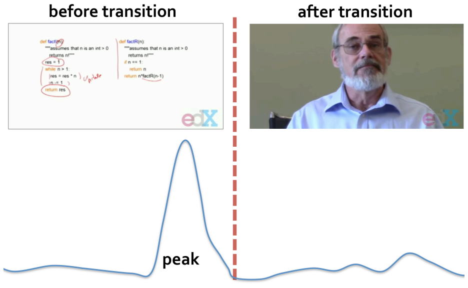

Figure 1. An example interaction peak near a scene change in a lecture video.

minutes long), informal (professors at a desk rather than behind a podium), and web-optimized (rather than in-class lecture captures) videos lead to higher engagement [2]. These analyses have implications for video authoring, editing, and interface design, and provide a richer understanding of video learning on MOOCs.

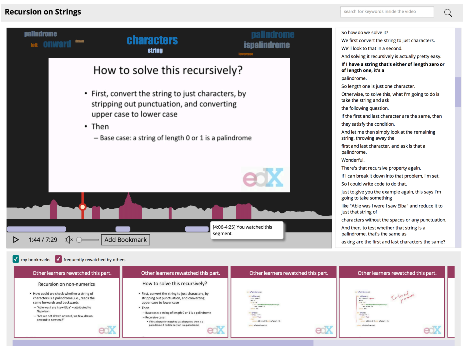

Figure 2. LectureScape: lecture video player powered by interaction data.

Video interface that evolves with data. Adapting the video presentation to the "online classroom" can directly improve the video learning experience. **LectureScape** (Figure 2) [3] is an enhanced video player for educational content online, powered by data on learners' collective video watching behavior. LectureScape dynamically adapts to thousands of learners' interaction patterns to make it easier to rewatch, skim, search, and review. By analyzing the viewing data as well as the content itself, LectureScape introduces a set of data-driven interaction techniques that augment existing video interface widgets: a 2D video timeline with an embedded visualization of collective navigation traces; dynamic and non-linear timeline dragging; data-enhanced transcript search and keyword summary; automatic display of relevant still frames next to the video; and a visual summary representing points with high learner activity. Participants in a user study commented that "it feels like watching with other students" and it was "more classroom-y" to watch videos with LectureScape, which shows how large-scale interaction data can support social and interactive video learning. This project has been published at UIST 2014 [3], a top-tier venue in HCI, and was featured by over 30 media outlets, including Forbes, eCampusNews, BostonInno, and MIT News Office. I am currently collaborating with edX to integrate LectureScape into the edX platform, and will soon release it as open source.

Adapting interaction behavior to user interest. LectureScape demonstrates an example of how signals of user interest can be used to dynamically improve the user interface. I extended this idea to a scrolling technique for touchscreen devices. The **content-aware kinetic scrolling** (CAKS) technique [4] uses social signals from web page content (e.g., number of likes, shares, or recommendations on social media) to model collective user interest. Based on the degree of interest model, CAKS dynamically modifies the scrolling behavior by applying additional friction around points of high interest within the page. This draws user attention to interesting content without cluttering the limited visual space.

Active Learnersourcing: Learner prompts contribute to new learning materials

How-to videos contain worked examples and step-by-step instructions for how to complete a task (e.g., math, cooking, programming, graphic design). My formative study demonstrated the navigational, self-efficacy, and performance benefits of having step-by-step information about the solution [5]. Education research has shown powerful learning gains in presenting the solution structure and labels for groups of steps (subgoals) to learners. However, such information is not available for most existing how-to videos online, and requires substantial expert efforts to collect. I have created scalable methods for extracting steps and subgoals from existing videos that do not require experts, as well as an alternative video player where the solution structure is displayed alongside the video. These techniques actively prompt learners to contribute structured information in an in-video quiz format.

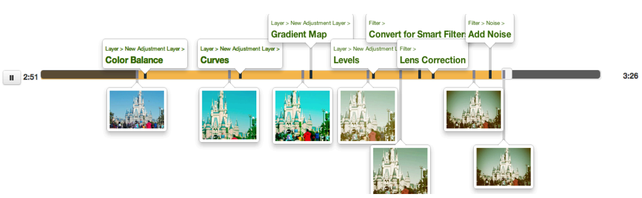

Figure 3. ToolScape: Step-aware tutorial video player.

Extracting step-by-step information from how-to videos. Based on findings from the formative study, I built **ToolScape**, a video player that displays step descriptions and intermediate result thumbnails in the video timeline (Figure 3) [5]. To enable non-experts to successfully extract step-by-step structure from existing how-to videos at scale, I designed a three-stage crowdsourcing workflow. It applies temporal clustering, text processing, and visual analysis algorithms to merge crowd output. The workflow successfully annotated 75 cooking, makeup, and Photoshop videos on YouTube of varying styles, with a quality comparable to trained annotators across all domains.

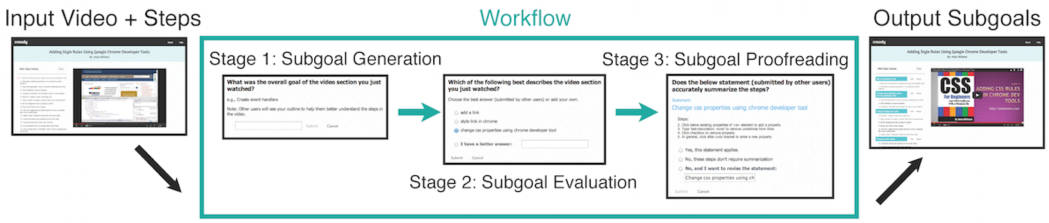

Figure 4. Crowdy: Learnersourcing workflow for summarizing steps in a how-to video.

Learnersourcing section summaries from how-to videos. Taking a step further, we asked if learners, both an intrinsically motivated and uncompensated crowd, can generate summaries of individual steps at scale. This research question resulted in a learnersourcing workflow that periodically prompts learners who are watching the video to answer one of the pre-populated questions, such as "what was the overall goal of the video section you just watched?" (Figure 4) [6]. The system dynamically determines which question to display depending on how much information has already been gathered for that section in the video, and the questions are designed to engage learners to reflect on the content. Learners' answers help generate, evaluate, and proofread subgoal labels, so that future learners can navigate the video with the solution summary. We deployed **Crowdy**, a live website with the learnersourcing workflow implemented on a set of introductory web programming videos. The 25-day deployment attracted more than 1,200 learners who contributed hundreds of subgoal labels and votes. A majority of learner-generated subgoals were comparable in quality to expert-generated ones, and learners commented that the system helped them grasp the material.

When combined, the two methods (ToolScape for individual steps and Crowdy for subgoal labels) can fully extract a hierarchical solution structure from existing how-to videos. I am currently conducting a controlled study to evaluate the pedagogical benefits of participating in our learnersourcing workflow and interacting with learnersourced summaries.

Designing systems for communities and crowds that care

I have also explored the domains of collaborative planning and civic engagement, in which a community of users perform inherently meaningful tasks while collaborating in the sensemaking and content generation process.

Room/ Time	Blue	Bordeaux	252B	352AB	Havane	241
Mon 11:00-12:20	+	Text Visualization 4	Call All Game Changers: 1	Novel Programmin 5	Lifetime Research Award	Fabrication 6 1
Mon 14:00-15:20	Language 1	Place meets Engagemen 1	Will Massive Online Open 2	+	Enterprise and online communitie 1	Evaluation Methods 1 2
Mon 16:00-17:20	Managment of Knowledge 1	Automated Usability / Evaluation 3 1	Theory and Practice in UX	Table and Floors	Smart Tools for Smart Work 2	Large and public Displays 1 2

Figure 5. Cobi: Conference scheduling tool powered by community input.

Community-driven conference scheduling. The onus of large-scale event planning often falls on a few organizers, and it is difficult to reflect the diverse needs and desires of community members. The **Cobi** project engages an entire academic community in planning a large conference. Cobi elicits community members' preferences and constraints, and provides a scheduling tool that empowers organizers to take informed actions toward improving the schedule (Figure 5) [7]. Community members' self-motivated interactions and inputs guide the conflict resolution and schedule improvements. Because each group within a community has different information needs, motivations, and interests in the schedule, we designed custom applications for different groups: Frenzy [8] for program committee members to group and label papers sharing a common theme; authorsourcing [7, 9] for paper authors to indicate papers relevant to theirs; and Confer [10] for attendees to bookmark papers of interest. Cobi has scheduled CHI and CSCW, two of the largest conferences in HCI, since 2013. It has successfully resolved conflicts and incorporated preferences in the schedule, with input from hundreds of committee members, and thousands of authors and attendees. I was the main designer and developer for the authorsourcing and scheduling applications [7], and helped the design of Frenzy [8] and Confer [10].

Budgetary discussion support for taxpayers. The extensiveness and complexity of a government budget hinder taxpayers from understanding budgetary information and participating in deliberation. In collaboration with economists, we built interactive and collaborative web platforms in which users can contribute to the overall "pulse" of the public's understanding and sentiment about budgetary issues. **Factful** [11] is an annotative news reading application that enhances the article with fact-checking support and contextual budgetary information. Users' fact-checking requests and results are accumulated to help future users engage in fact-oriented discussions.

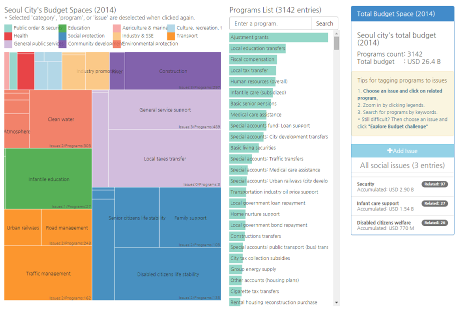

Figure 6. BudgetMap: Issue-driven visual exploration of government budgets.

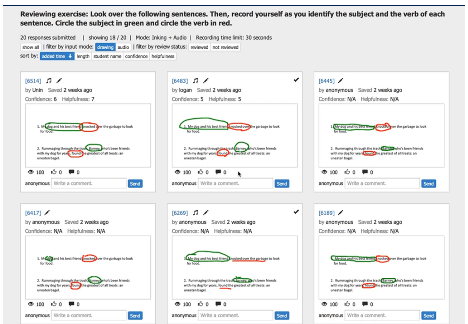

Figure 7. RIMES: Interactive multimedia exercises in lecture videos.

BudgetMap (Figure 6) [12] allows users to navigate the government budget with social issues of their interest. Users can make links between government programs and social issues by tagging. In our 5-day live deployment, more than 1,600 users visited the site and made 697 links between social issues and budget programs. The government of South Chungcheong Province in South Korea has decided to officially adopt BudgetMap.

Future Research Agenda

My research has introduced computational mechanisms in which users' lightweight contributions serve a bigger cause: learners improve content and interfaces for future learners, paper authors and attendees help with conference scheduling, and taxpayers make fact checking requests and link social issues to budget items. My research will continue to lower the barrier to individuals' participation and impact within a community, by engaging them in activities meaningful to both themselves and the community.

Innovative learnersourcing applications. In the near future, I plan to push the boundaries of learnersourcing by broadening the application scope and domain. As part of a potential direction, I built a new system, RIMES (Figure 7), for easily authoring, recording, and reviewing interactive multimedia exercises embedded in lecture videos [13]. With RIMES, teachers can prompt learners to record their responses to an activity using video, audio, and inking while watching lecture videos. Teachers can then review and interact with all the learners' responses in an aggregated gallery. I plan to extend RIMES to learnersource alternative self-explanations within a video, which can enable multiple learning paths and personalization for learners. A technical challenge will be in interpreting and labeling multimedia inputs from learners. Another avenue for future research is supporting diverse problem domains: a programming environment that learnersources code snippets and documentation; a writing tool that learnersources alternate expressions and phrases for inspiration and learning; and a graphical design tool that learnersources visual assets other learners can extend.

Learning platforms of the future. While existing online learning platforms have provided access to more learners, the learning experience is still limited to passively watching videos and answering canned questions. I envision educational technologies that truly scale: a course that is created, organized, and taught entirely by learners. Learners will actively engage in 1) creating various artifacts including quizzes, explanations, and examples, 2) providing feedback and iterating on the artifacts, and 3) labeling the artifacts with metadata for better search and organization. With a learnersourced course that builds itself, learners are at the center of both the instructional design and learning experience.

Science of learning at scale. Despite the rapid growth, we do not yet have good answers to whether, what, how, and why people learn in massive learning platforms. The field needs to develop theories, instructional guidelines, and design principles for learning at scale settings. In collaboration with education researchers, I plan to build formal and data-driven frameworks for conducting controlled experiments and adaptively modifying the learning interface. The frameworks can help researchers experiment with various learnersourcing models in vivo, with different video formats, interaction patterns, learner prompts, and adaptation techniques. These efforts will spur the design of new instructional formats and technologies, which in turn can advance the emerging science of learning at scale.

Sociotechnical application design. My research program has uniquely applied insights from crowdsourcing (microtasks and workflow design) to lower the barrier to participation for individuals within a community. I believe this framework has potential for broader societal impact. I plan to design novel coordination and incentive mechanisms for broader domains including civic engagement, healthcare, and accessibility. I will investigate generalizable design principles and computational methods applicable across multiple domains. Furthermore, I am interested in capturing community interactions beyond clickstream and answers to given prompts. How can we capture discussion, collaboration, decision making, and creative processes by individuals and groups at scale? I will continue to build systems that are used by real people in answering the proposed research questions.

Collaboration. In graduate school, I have been fortunate to work with over 60 collaborators from over 20 different institutions, spanning a variety of domains including machine learning, natural language processing, computer vision, artificial intelligence, systems, communications, education, psychology, and economics. In a new environment, I look forward to aiming for more ambitious research goals and impact with my collaborators.

References

[1] **Juho Kim**, Philip J. Guo, Daniel T. Seaton, Piotr Mitros, Krzysztof Z. Gajos, Robert C. Miller. Understanding In-Video Dropouts and Interaction Peaks in Online Lecture Videos. In Proc. of *L@S 2014: ACM Conference on Learning at Scale.*

[2] Philip J. Guo, **Juho Kim**, Rob Rubin. How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In Proc. of *L@S 2014: ACM Conference on Learning at Scale.*

[3] **Juho Kim**, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, Robert C. Miller. Data-Driven Interaction Techniques for Improving Navigation of Educational Videos. In Proc. of *UIST 2014: ACM Symposium on User Interface Software and Technology.*

[4] **Juho Kim**, Amy X. Zhang, Jihee Kim, Robert C. Miller, Krzysztof Z. Gajos. Content-Aware Kinetic Scrolling for Supporting Web Page Navigation. In Proc. of *UIST 2014: ACM Symposium on User Interface Software and Technology.*

[5] **Juho Kim**, Phu Nguyen, Sarah Weir, Philip J. Guo, Robert C. Miller, Krzysztof Z. Gajos. Crowdsourcing Step-by-Step Information Extraction to Enhance Existing How-to Videos. In Proc. of *CHI 2014: ACM Conference on Human Factors in Computing Systems.* **Best of CHI Honorable Mention.**

[6] Sarah Weir, **Juho Kim**, Krzysztof Z. Gajos, Robert C. Miller. Learnersourcing Subgoal Labels for How-to Videos. In Proc. of *CSCW 2015: ACM Conference on Computer-Supported Cooperative Work and Social Computing, to appear.*

[7] **Juho Kim**, Haoqi Zhang, Paul André, Lydia B. Chilton, Wendy Mackay, Michel Beaudouin-Lafon, Robert C. Miller, Steven P. Dow. Cobi: A Community-Informed Conference Scheduling Tool. In Proc. of *UIST 2013: ACM Symposium on User Interface Software and Technology.*

[8] Lydia Chilton, **Juho Kim**, Paul André, Felicia Cordeiro, James Landay, Dan Weld, Steven P. Dow, Robert C. Miller, Haoqi Zhang. Frenzy: Collaborative Data Organization for Creating Conference Sessions. In Proc. of *CHI 2014: ACM Conference on Human Factors in Computing Systems*. **Best of CHI Honorable Mention.**

[9] Paul André, Haoqi Zhang, **Juho Kim**, Lydia B. Chilton, Steven P. Dow, Robert C. Miller. Community Clustering: Leveraging an Academic Crowd to Form Coherent Conference Sessions. In Proc. of *HCOMP 2013: AAAI Conference on Human Computation and Crowdsourcing.* **Notable Paper Award.**

[10] Anant Bhardwaj, **Juho Kim**, Steven P. Dow, David Karger, Sam Madden, Robert C. Miller, Haoqi Zhang. Attendee-sourcing: Exploring the Design Space of Community-Informed Conference Scheduling. In Proc. of *HCOMP 2014: AAAI Conference on Human Computation and Crowdsourcing.*

[11] **Juho Kim**, Eun-Young Ko, Jonghyuk Jung, Chang Won Lee, Nam Wook Kim, Jihee Kim. Factful: Engaging Taxpayers in the Public Discussion of a Government Budget. In submission to *CHI 2015: ACM Conference on Human Factors in Computing Systems.* (under review)

[12] Nam Wook Kim, Chang Won Lee, Jonghyuk Jung, Eun-Young Ko, **Juho Kim**, Jihee Kim. BudgetMap: Supporting Issue-Driven Navigation for Government Budget. In submission to *CHI 2015: ACM Conference on Human Factors in Computing Systems.* (under review)

[13] **Juho Kim**, Elena L. Glassman, Andrés Monroy-Hernández, Meredith Ringel Morris. RIMES: Embedding Interactive Multimedia Exercises in Lecture Videos. In submission to *CHI 2015: ACM Conference on Human Factors in Computing Systems.* (under review)