Organic Crowdsourcing Systems

Juho Kim
Stanford University and KAIST
juho@juhokim.com

Abstract

Crowdsourcing has the potential to enable distributed teamwork at scale. While some existing crowdsourcing systems have attracted voluntary crowds to achieve shared goals, little research has investigated underlying design principles and mechanisms that make these systems successful. I explore *organic crowdsourcing*, a form of crowdsourcing in which people collectively contribute work while engaging in a meaningful experience themselves. Two major properties of organic crowdsourcing are: (1) people individually benefit by participating in the crowdsourcing workflow, and (2) the collectively produced outcome is of value to the crowd. I present organic crowdsourcing applications from three domains: learning from an instructional video, planning an academic conference, and understanding a government budget. In these domains, the crowd is intrinsically motivated to participate in the crowd work, while the system prompts, collects, processes, and presents the crowd's collective contributions to produce a valuable outcome for the users of the same system.

The advances in remote communication technologies and coordination methods have enabled support for various forms of distributed human teamwork. On the one hand, there are small teams with a clear organizational structure and sense of membership: e.g., a 10-person team inside a company in which members work in remote locations. On the other hand, there are large groups of people online who work on some shared task without a clear organizational structure and sense of membership: e.g., crowd workers recruited to work on a task on Amazon Mechanical Turk. While the latter may not be seen as teamwork in a strict sense, more crowdsourcing applications are incorporating teamwork structure and shared goal-setting to enable distributed teamwork at scale. Some major characteristics of crowdsourcing systems are: (1) a large number of users make micro-contributions, (2) social interaction between users tends to be indirect or implicit by design, and (3) work between users is loosely coupled or independent. While crowdsourcing has enabled tackling problems that humans or machines cannot solve alone, systems supporting large-scale, crowdsourced work face major challenges: quality control, limited complexity of produced work, and lack of workers' intrinsic motivation.

To overcome these limitations, research in crowdsourcing has introduced organizational structure (Retelny et al. 2014), dynamic coordination (Kulkarni, Can, and Hartmann 2012; Kittur et al. 2011), and design patterns and workflows (Bernstein et al. 2010; Little et al. 2009). There have also been crowdsourcing systems with intrinsically motivated users who play a game for fun (von Ahn and Dabbish 2004; von Ahn et al. 2008), contribute to science [1], or provide service to a community [2]. In my research, I explore *organic crowdsourcing*, a form of crowdsourcing in which people collectively contribute work while engaging in a meaningful experience themselves. Two major properties of organic crowdsourcing are: (1) people individually benefit by participating in the crowdsourcing workflow, and (2) the collectively produced outcome is of value to the crowd.

To better illustrate what organic crowdsourcing is, let's compare organic crowdsourcing against other crowdsourcing models. A fundamental difference between organic crowdsourcing and paid crowdsourcing (e.g., Amazon Mechanical Turk, Upwork) is the incentive structure. While organic crowdsourcing does not pay for people's work, it is important to note that people's time and effort are the primary cost in organic crowdroucing tasks. Organic crowdsourcing also differs from human computation systems such as the ESP Game (von Ahn and Dabbish 2004) or ReCaptcha (von Ahn et al. 2008). In the ESP Game, a user benefits by playing a game for fun, but the produced outcome is labels for images used to improve image search but not to directly benefit the user. Similarly, in ReCaptcha, a user benefits by proceeding to their main task (e.g., logging in to a website) after verifying that they are human, but the produced outcome is digitized text that does not directly benefit the user.

While some existing organic crowdsourcing systems have attracted a large crowd to participate voluntarily, what makes some of them successful and not others remains a black box. To promote the design of future organic crowdsourcing systems, it is important to investigate generalizable design guidelines and principles, design coordination mech-

[1] https://www.zooniverse.org/
[2] http://www.archives.gov/citizen-archivist/

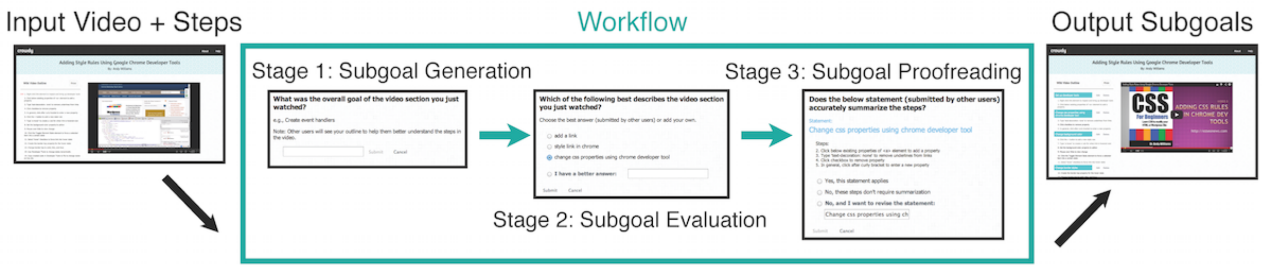

Figure 1: Crowdy is a video player that coordinates learners' collective efforts to generate subgoal labels, which are summarized, abstract descriptions for a group of low-level steps in a procedural task.

anisms, and crowdsourcing workflows. In this position paper, I present organic crowdsourcing applications from three domains: learning from an instructional video, planning an academic conference, and understanding a government budget. In these domains, a crowd's intrinsic motivation is to learn new knowledge and master skills, to attend interesting and relevant talks at a conference, and to make sense of how the government spends their tax money, respectively. These systems prompt, collect, process, and present the crowd's collective contributions to produce a valuable outcome for the users of the same system.

Learnersourcing: Improving Learning with Collective Learner Activity

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? My primary approach has been learnersourcing (Kim 2015), in which learners collectively generate novel content and interfaces for future learners while engaging in a meaningful learning experience themselves. My research demonstrates that interfaces powered by learnersourcing can enhance content navigation, create a sense of learning with others, and ultimately improve learning.

How-to videos contain worked examples and step-by-step instructions for how to complete a task (e.g., math, cooking, programming, graphic design). Our formative study demonstrated the navigational, self-efficacy, and performance benefits of having step-by-step information about the solution (Kim et al. 2014). 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. We 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.

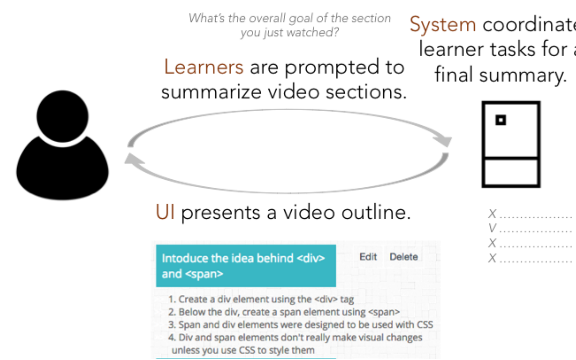

Figure 2: As learners watch a video, Crowdy pauses the video and prompts learners to summarize video sections. These tasks are designed to improve learners' understanding of the material. At the same time, the system coordinates learner tasks to reach a final set of summary labels for the clip. Crowdy dynamically displays the collectively generated video outline as future learners visit and watch the video.

Extracting step-by-step information from how-to videos: To enable non-experts to successfully extract step-by-step structure from existing how-to videos at scale, we designed a three-stage crowdsourcing workflow (Kim et al. 2014). 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.

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 2) (Weir et al. 2015). 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 learn-

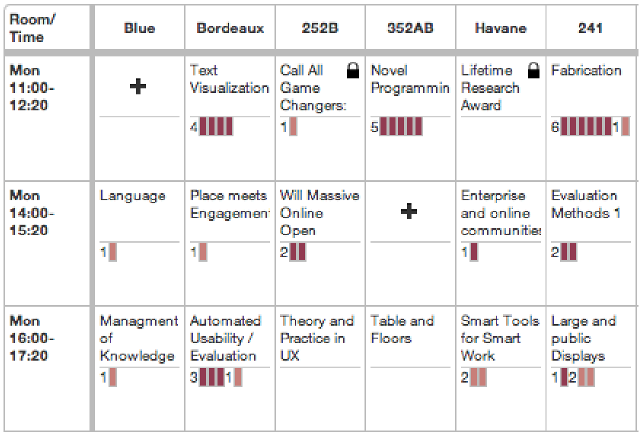

Figure 3: Cobi is a conference scheduling tool powered by community input. Community members indicate their preferences and constraints in the schedule with the expectation of either presenting their papers in a relevant session or seeing interesting papers. Conference organizers use this information to make informed schedule-related decisions.

ers 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 and statistics videos. A study with Turkers showed higher learning outcomes with Crowdy when compared against the baseline video interface. The Crowdy group also performed as well as the group that were shown expert-generated labels. For the four videos with the highest participation, we found that a majority of learner-generated subgoals were comparable in quality to expert-generated ones. Learners commented that the system helped them grasp the material, suggesting that our workflow did not detract from the learning experience.

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 3) (Kim et al. 2013). 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 (Chilton et al. 2014) for program committee members to group and label papers sharing a common theme; authorsourcing (Kim et al. 2013; André et al. 2013) for paper authors to indicate papers relevant to theirs; and Confer (Bhardwaj et al. 2014) 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.

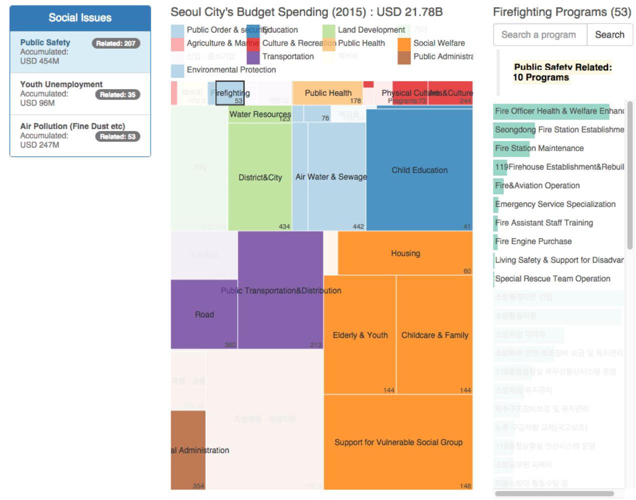

Figure 4: BudgetMap engages taxpayers to provide social issue tags for budget items. While taxpayers' awareness and understanding of budgetary issues increases, the interface enables issue-driven budget navigation.

Improving Sensemaking and Awareness of Government Budget

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 (Kim et al. 2015) 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. BudgetMap (Figure 4) (Kim et al. 2016) 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. Our evaluation shows that participants' awareness and understanding of budgetary issues increased after using BudgetMap, while they collaboratively identified issue-budget links with quality comparable to expert-generated links.

Vision

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. Organic crowdsourcing systems can help lower the barrier to individuals' participation and impact within a community, by engaging them in activities meaningful to both themselves and the community. I believe this framework has potential for broader societal impact in designing sociotechnical applications. Some future research directions include: (1) designing novel coordination and incentive mechanisms for broader domains including civic engagement, healthcare, and accessibility; (2) identifying generalizable design principles and computational methods applicable across multiple domains; and (3) capturing rich, contextual community interactions such as discussion, collaboration, decision making, and creative processes by individuals and groups at scale.

Acknowledgement

I would like to thank Rob Miller and Krzysztof Gajos for their thoughtful comments on all the projects and vision presented in this paper.

References

André, P.; Zhang, H.; Kim, J.; Chilton, L.; Dow, S. P.; and Miller, R. C. 2013. Community clustering: Leveraging an academic crowd to form coherent conference sessions. In *First AAAI Conference on Human Computation and Crowdsourcing*, HCOMP '13.

Bernstein, M. S.; Little, G.; Miller, R. C.; Hartmann, B.; Ackerman, M. S.; Karger, D. R.; Crowell, D.; and Panovich, K. 2010. Soylent: A word processor with a crowd inside. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, 313–322. New York, NY, USA: ACM.

Bhardwaj, A.; Kim, J.; Dow, S.; Karger, D.; Madden, S.; Miller, R.; and Zhang, H. 2014. Attendee-sourcing: Exploring the design space of community-informed conference scheduling. In *Second AAAI Conference on Human Computation and Crowdsourcing*.

Chilton, L. B.; Kim, J.; André, P.; Cordeiro, F.; Landay, J. A.; Weld, D. S.; Dow, S. P.; Miller, R. C.; and Zhang, H. 2014. Frenzy: Collaborative data organization for creating conference sessions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, 1255–1264. New York, NY, USA: ACM.

Kim, J.; Zhang, H.; André, P.; Chilton, L. B.; Mackay, W.; Beaudouin-Lafon, M.; Miller, R. C.; and Dow, S. P. 2013. Cobi: A community-informed conference scheduling tool. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, 173–182. New York, NY, USA: ACM.

Kim, J.; Nguyen, P. T.; Weir, S.; Guo, P. J.; Miller, R. C.; and Gajos, K. Z. 2014. Crowdsourcing step-by-step information extraction to enhance existing how-to videos. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, 4017–4026. New York, NY, USA: ACM.

Kim, J.; Ko, E.-Y.; Jung, J.; Lee, C. W.; Kim, N. W.; and Kim, J. 2015. Factful: Engaging taxpayers in the public discussion of a government budget. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, 2843–2852. New York, NY, USA: ACM.

Kim, N. W.; Jung, J.; Ko, E.-Y.; Han, S.; Lee, C. W.; Kim, J.; and Kim, J. 2016. Budgetmap: Engaging taxpayers in the issue-driven classification of a government budget. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work & Social Computing, to appear*, CSCW '16.

Kim, J. 2015. *Learnersourcing: Improving Learning with Collective Learner Activity*. Ph.D. Dissertation, MIT, 77 Massachusettes Avenue, Cambridge, MA 02139.

Kittur, A.; Smus, B.; Khamkar, S.; and Kraut, R. E. 2011. Crowdforge: Crowdsourcing complex work. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, 43–52. New York, NY, USA: ACM.

Kulkarni, A.; Can, M.; and Hartmann, B. 2012. Collaboratively crowdsourcing workflows with turkomatic. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, 1003–1012. New York, NY, USA: ACM.

Little, G.; Chilton, L. B.; Goldman, M.; and Miller, R. C. 2009. Turkit: Tools for iterative tasks on mechanical turk. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, HCOMP '09, 29–30. New York, NY, USA: ACM.

Retelny, D.; Robaszkiewicz, S.; To, A.; Lasecki, W. S.; Patel, J.; Rahmati, N.; Doshi, T.; Valentine, M.; and Bernstein, M. S. 2014. Expert crowdsourcing with flash teams. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, 75–85. New York, NY, USA: ACM.

von Ahn, L., and Dabbish, L. 2004. Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, 319–326. New York, NY, USA: ACM.

von Ahn, L.; Maurer, B.; McMillen, C.; Abraham, D.; and Blum, M. 2008. recaptcha: Human-based character recognition via web security measures. *Science* 321(5895):1465–1468.

Weir, S.; Kim, J.; Gajos, K. Z.; and Miller, R. C. 2015. Learnersourcing subgoal labels for how-to videos. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, 405–416. New York, NY, USA: ACM.