

# Programming the crowds

October 27 2010, by Larry Hardesty, MIT News

---



Graphic: Christine Daniloff

At the Association for Computing Machinery's 23rd symposium on User Interface Software and Technology in October, members of the User Interface Design Group at MIT's Computer Science and Artificial Intelligence Laboratory walked off with awards for both best paper and best student paper.

Both papers describe [software](#) that uses Amazon's Mechanical Turk service to farm tasks out to human beings sitting at Internet-connected computers. The best-paper prize went to [VizWiz](#), a cell-phone application that enables blind people to snap photographs and, within a minute or so, receive audible descriptions of the objects depicted. The

best-student-paper prize went to Soylent, a program that distributes responsibility for editing text to hundreds of people in such a way that highly reliable results can be culled in as little as 20 minutes.

Launched in 2005, [Mechanical Turk](#) is an Internet marketplace, where so-called requesters can upload data and offer payment for the completion of simple tasks that computers can't perform reliably. A requester, for instance, might split an audio file into 30-second chunks and offer five cents for the transcription of each one. But Mechanical Turk has a few drawbacks. One is its fairly clunky interface. Others are the difficulty of getting results in real time and of getting reliable results to complex tasks. The MIT researchers' papers address all three problems.

## People in the program

Both VizWiz and Soylent operate as free-standing applications: The user wouldn't necessarily know that they use Mechanical Turk at all, only that they take somewhat longer to return results than computer programs generally do. VizWiz, whose design was led by Jeffrey Bigham, an assistant professor at the University of Rochester who was a visiting professor at CSAIL last fall, uses a variety of tricks to get results within seconds. One of these is to recruit Mechanical Turk workers — or “turkers” — as soon as the application launches, on the assumption that it will imminently provide data for analysis. To hold the recruited workers' attention, and to prime them for the type of task they'll be asked to perform, the VizWiz system gives them a series of already-solved problems to work on, until the cell-phone application has produced a new photo for identification.

“The use scenario is just fantastic,” says Ben Bederson, an associate professor at the University of Maryland's Human-Computer Interaction Lab, who was not involved in the project. “It's not only visually impaired

people. I was traveling in China, and I can't read the road signs. There are a number of times where people want real-time help." Bederson adds that he considers the most innovative aspect of the program the technique it uses to enable blind people to take photos in the first place. "They came up with this mechanism where they used real-time computer vision and sound generation to give audio feedback of whether or not you're pointing at the right thing," he says. "That's a totally new concept. That's incredible."

Nonetheless, Bederson says, Soylent may be the more innovative of the two systems. "The nature of this labor force is that it's very disconnected," Bederson says. "There's a huge amount of lack of trust and cheating." For instance, Bederson says, his own research group tried to use Mechanical Turk to assign two-paragraph movie reviews a simple binary rating: favorable or unfavorable. "Most of the answers were basically garbage," he says. A turker who submitted a random rating was paid as much as a turker who actually took a minute to read the review. "There's been a lot of work on anti-cheating mechanisms," Bederson says, but Soylent "goes a step farther in coming up with mechanisms to engage work that before this paper people didn't think was possible to do with this kind of labor force."

## **Between extremes**

In the MIT researchers' experiments, Soylent recruited turkers to perform two different tasks: one was to copyedit a document of roughly seven paragraphs; the other was to shorten a document. "The naïve thing you would think to do is put your paragraph on Mechanical Turk and say, 'Hey, could you please make this shorter?'" says Michael Bernstein, a PhD student in the Computer Science and Artificial Intelligence Laboratory who together with associate professor Rob Miller, who heads the User Interface Design Group, led the Soylent project. "A huge set of workers will simply find the simplest single thing they can do, make one

short edit somewhere near the beginning of your document, and run away with your money.” Another group, whom Bernstein and Miller call “eager beavers,” will perform all kinds of unnecessary edits that may even introduce errors. “They’re trying to help, but they’re doing more harm than good,” Bernstein says.

So the MIT researchers, and their colleagues at the University of California, Berkeley, and the University of Michigan, split editing tasks into three stages, which they label Find, Fix, and Verify. Initially, turkers are recruited simply to identify sections of a text that are wordy or contain grammatical errors. Those sections are then abstracted from the text and passed on to additional turkers who propose solutions. And finally, another round of turkers evaluate the proposed solutions, weeding out those that are ungrammatical. In experiments, the researchers found that \$1.50 per paragraph would elicit good results within 20 minutes; the cost would go down to about 30 cents per paragraph if the user was willing to wait a couple hours. Either way, by the time the results came back, literally hundreds of turkers would have worked on the document.

“There’s a lot of work nowadays on Mechanical Turk,” says Luis von Ahn, an assistant professor of computer science at Carnegie Mellon University. “I find that most of it is pretty crappy. But this is one of the only papers I saw that told me something that I didn’t actually know.” Von Ahn is famous as the inventor of captchas, those online tests that ask the user to re-type a line of distorted text in order to verify that he or she is not a spam robot. According to von Ahn, work like Soylent and VizWiz is valuable for revealing what’s possible with Internet-coordinated, distributed labor. But, he says, he would like to see researchers develop a better theoretical understanding of the relationship between the size of tasks, cost and time, which would explain why systems like Soylent work where others don’t. Given a problem, von Ahn asks, “How do you split it up into the right-size pieces so that at the end

of the day you get the most people doing it and probably the cheapest?” In some instances, he says, “A 20-minute task might be unmanageable, but a nine-minute task is just right. It would be really nice to have more of a theory of this.”

---

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Programming the crowds (2010, October 27) retrieved 9 April 2024 from <https://phys.org/news/2010-10-crowds.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--