

# Crowd workers are not online Shakespeares, but Carnegie Mellon research shows they can write

February 2 2011

---

Writing can be a solitary, intellectual pursuit, but researchers at Carnegie Mellon University have shown that the task of writing an informational article also can be accomplished by dozens of people working independently online.

Each person in the CMU experiments completed just a sliver of the work of preparing an article, such as preparing an outline, gathering facts or assembling facts into simple prose. The "authors" never even spoke with each other. But the research team led by Aniket Kittur, assistant professor in CMU's Human-Computer Interaction Institute (HCII), found that the crowdsourced articles compared favorably with articles written by a single author and with Simple English Wikipedia entries.

"This is exciting because collaborative crowdsourcing could change the future of work," Kittur said. "We foresee a day when it will be possible to tap into hundreds of thousands or millions of workers around the globe to accomplish creative work on an unprecedented scale."

Kittur, along with Robert Kraut, professor of human-computer interaction, and Boris Smus, a student in HCII's joint master's degree program with the University of Madeira, have created a framework called CrowdForge that breaks down complex tasks into simple, independent micro-tasks that can be completed rapidly and cheaply. Their technical paper is available online at <http://reports->

[archive.adm.cs.cmu.edu/anon/hcii/abstracts/11-100.html](http://archive.adm.cs.cmu.edu/anon/hcii/abstracts/11-100.html).

Jim Giles and MacGregor Campbell, San Francisco-based science journalists, have created a blog, [www.mybossisarobot.com](http://www.mybossisarobot.com), that will explore the use of CrowdForge for preparing science news articles based on research reports.

Crowdsourcing has become a powerful mechanism for accomplishing work online. Millions of volunteers have performed tasks such as cataloging Martian landforms (<http://beamartian.jpl.nasa.gov>) and translating text into machine-readable form (<http://recaptcha.com>).

In the Carnegie Mellon experiments, crowdsourced work was performed through Amazon's Mechanical Turk (MTurk), an online marketplace for work. Employers can post simple, self-contained tasks on MTurk that workers, or "turkers," complete in return for a small fee, usually a few cents. Typical tasks include identifying objects in photos, writing product descriptions and transcribing audio recordings.

"But much of the work required by real-world organizations requires more time, cognitive effort and coordination among co-workers than is typical of these crowdsourcing efforts," Kittur said. Most turkers, for instance, refuse long, complex tasks because they are paid so little in return.

To accomplish these complex tasks, the CMU researchers approached the crowdsourcing market as if it was a distributed computing system, like the large computer systems used for Web searches. In a distributed computing system, computations are divided up in such a way that smaller chunks can be solved simultaneously by large numbers of processors and failures by individual processors won't undermine the entire process. Google, for instance, uses a framework called MapReduce in which queries are divided, or mapped, into sub-problems

that can be solved simultaneously by numerous computers. The results of the computations then are combined, or reduced, to answer the query.

The framework developed by the CMU researchers, called CrowdForge, likewise divides up complex tasks so that many individuals can complete parts of the overall task and then provides a means for coordinating, combining and evaluating their work.

To prepare a brief encyclopedia article, for instance, CrowdForge would assign several people the task of writing an outline; as a quality control measure, a second set of workers might be tasked with voting for the best outline, or combining the best parts of each outline into a master outline. Subsequent sub-tasks might include collecting one fact for a topic in the outline. Finally, a worker might be given the task of taking several of the facts collected for a topic and turning them into a paragraph, or combining several paragraphs in proper order for an article.

In preparing five such articles on New York City, this method required an average of 36 sub-tasks for each article, at an average cost of \$3.26. The articles averaged 658 words. The researchers then paid eight individuals \$3.05 each to produce short articles on the same subjects; the average length was 393 words. When 15 people compared the articles, they rated the group-written articles of higher quality than those produced by individuals and about the same as a [Wikipedia](#) entry on the topic. The variability — the range from the best to the worst article — was lower for the crowdsourced articles.

"We were surprised at how well CrowdForge worked," Kittur said.

"Admittedly, none of these articles is going to win any awards. But the ratings weren't bad considering that the work of dozens of people had to be coordinated to produce these pieces."

Kittur said the significance of CrowdForge is that it shows crowdsourcing of creative work is feasible, not that it can drive down the cost of articles. "We used MTurk as a source of workers," he noted, "but other users might tap into writers and researchers within an organization or into an existing network of freelancers."

Provided by Carnegie Mellon University

Citation: Crowd workers are not online Shakespeares, but Carnegie Mellon research shows they can write (2011, February 2) retrieved 26 June 2024 from <https://phys.org/news/2011-02-crowd-workers-online-shakespeares-carnegie.html>

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.