# A New View on HCI and Crowdsourcing

**James A. Landay**
Computer Science & Engineering
DUB Institute, University of Washington
Seattle, WA 98195-2350
**landay@cs.washington.edu**

**ABSTRACT**
There are two important uses of crowdsourcing we should be exploring. One is to focus on the partnership between human computation and machine computation. We should allow humans to contribute data that will eventually allow a machine to carry out a task, rather than letting the human judgments be the ultimate goal. The other important use is to improve end-user programming applications. Human computation is often used to fill the gaps where machines have fallen short. Many end-user programming systems are limited by what end-user affordances they can make. Instead of excluding those features from the system, allow the end-user to rely on a human to fill in the gaps.

**ACM Classification Keywords**
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## CROWDSOURCING EXPERIENCE

I have worked with graduate students at the University of Washington and Microsoft Research on two early stage projects in the area of Crowdsourcing. The first project, joint with Michael Toomim and Claus Pörtner, is an attempt to use crowdsourced systems to measure the economic utility of user interfaces and more generally social systems. In the second project, which is being carried out primarily by Lydia Chilton, we are trying to use crowdsourcing to help people who are learning a foreign language.

### Interaction Economics

Web site and the social-computational systems layered on top of them are complex phenomena that are clearly important to society. Yet we have a poor scientific understanding of them. A science depends on objective, empirical measurements, but we do not know how to measure and optimize the productive capacity of, for example, Wikipedia. A science uses experiments to verify hypotheses, but few researchers and designers can manipulate variables of web sites or social systems such as Facebook and Twitter, because they do not have access to the source code or the userbase. A science enables a community of scholars to replicate findings and generalize results, but the field of Human-Computer Interaction rarely replicates experiments, and we find it difficult to generalize knowledge from one system to another. We cannot even quantify the most fundamental resource of a web site – its ability to convince people to use it—because we do not know how to measure the utility of computer systems, and we do not know how to apply our science of utility (Economics) with our science of computer interfaces (Human-Computer Interaction).

Our work on Interaction Economics enables a more scientific approach to web design and social-computational systems. We develop methods that use Internet microtask labor markets, such as Amazon's Mechanical Turk [2], to create virtual scientific laboratories to experiment with web sites. These laboratories allow researchers and designers to easily, cheaply, and quickly test new interfaces and social arrangements with hundreds of real study participants. Furthermore, our methods allow a transformative way of looking at the objective of design. Rather than measure time-on-task, bug counts, or subjective survey reports of satisfaction, we enable designers to measure the quantitative degree to which users want to use their system. Borrowing from Economics, we call this the utility of a system.

Our transformative approach [4] merges the fields of Economics and HCI into a new field we call Interaction Economics. Interaction Economics quantifies user preference for different computer systems, modeling the utility differences that characterize their costs and benefits, and underlie user preference and choice. These methods will enable a new science that will let us experimentally design new systems, as well as optimize current systems, of greater complexity and novelty, with greater certainty of success. Our methodology relies on a novel use of Internet labor markets: a place where interactive tasks (HCI) meet incentives and markets (Economics). With controlled payments to real people we can prototype and study a wide range of web interfaces and social-computational phenomena that was not possible before.

### Crowdsourced Language Learning

A crucial element of language learning is the time learners spend listening and speaking their new language. Many learners at all skill levels become immersed in their new language: novice or intermediate students studying abroad for a semester, advanced learners who move for work or study to an area where their new language is primary, or even a learner in a classroom setting which emphasizes speaking and listening. Immersing environments force learners to use their skills but they can also be frustrating for both learners at all levels:

- For novice or intermediate learners, even though some experiences are repeated often, such as going to the same restaurant, the same supermarket, or listening to the same announcements on the subway, the experience goes by too quickly to absorb the new material.
- Intermediate learners often pick up enough of what is spoken to comprehend a native speaker, but would benefit from learning the things they did not understand.
- Advanced learners speak quite often and quite fluently to native speakers, however, they rarely receive correction even on common mistakes, from the native speakers because it is rude and interrupts the flow of conversation. This reinforces mistakes.
- Even in a classroom setting, students fail to pick up on all the vocabulary an instructor might use in attempt to replicate an immersive environment. There is not enough time to slow down the lesson and ask questions about everything.

Our solution to these problems is to capture audio recordings of immersive language situations, and from these actual experiences of learners, use human effort to create short lessons for the learner to use in review. Human effort is used for three things:

1. Translation/transcription abilities that computers currently lack
2. Common sense and cultural explanation of scenarios that computers also currently lack
3. The ability to understand individual user needs and the challenges of the environments they find themselves in and adapt lessons to them. This level of flexibility is not currently seen in computer-based learning programs.

## VISION FOR FUTURE CROWDSOURCING RESEARCH

There is a vast amount of interesting research in using crowdsourcing and human computation in novel systems and applications. One interesting direction is in using a Human Computation API in new systems. I am particularly interested in developing new architectures for the use of Human Computation as well as using Human Computation in new applications where the skills required of the "humans" is much more complex.

I believe an interesting future research direction is in defining an architecture for applications that partitions the computation between humans and machines in a more fluid way. For example, imagine a system that uses humans to carry out recognition tasks and in so doing labels underlying data that eventually a machine learning system starts to automatically take over when it can perform the tasks better or more cheaply, possibly using humans to simply check the work (e.g., ReCAPTCHA [1] or CardMunch [3]). How do we create an API that allows programmers to specify such a system, parameterize it appropriately, and then test it? These are fertile areas for new work.

The other area of human computation that I am especially interested in is in leveraging it for solving problems that are today considered much more skilled, e.g., for programming. Imagine an end-user programming system that allows non-programmers (e.g., user interface designers) to create much of an application, but when they reach a ceiling that current tools often break on they can instead use human computation to outsource a small piece of code that allows them to achieve their vision – all within the original end-user system. What types of tasks can be "out-sourced" in this way? How quickly can they be turned around so that the use of the original tool feels fluid? How much would you have to pay? How do you "test" the code and pick the most appropriate result? These are the research questions we will explore.

## BIOGRAPHY

James Landay is the Short-Dooley Professor of Computer Science & Engineering at the University of Washington, specializing in human-computer interaction. From 2003 through 2006 he was also the Laboratory Director of Intel Labs Seattle, a university affiliated research lab exploring ubiquitous computing. His current research interests include Automated Usability Evaluation, Demonstrational Interfaces, Mobile & Ubiquitous Computing, User Interface Design Tools, Crowdsourcing, and Web Design. He is spending his 2009-2011 sabbatical at Microsoft Research Asia in Beijing.

Landay received his B.S. in EECS from UC Berkeley in 1990 and M.S. and Ph.D. in CS from Carnegie Mellon University in 1993 and 1996, respectively. His Ph.D. dissertation was the first to demonstrate the use of sketching in user interface design tools. He was also the chief scientist and co-founder of NetRaker. In 1997 he joined the faculty in EECS at UC Berkeley, leaving as an Associate Professor in 2003.

## REFERENCES


1. Luis von Ahn, Ben Maurer, Colin McMillen, David Abraham and Manuel Blum. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. Science, September 12, 2008. pp 1465-1468.
2. Amazon's Mechanical Turk. http://www.mturk.com
3. Card Munch. http://www.cardmunch.com/
4. Toomim, M., T. Kriplean, C. Pörtner and J.A. Landay, Measuring Utility of Human-Computer Interactions (title may change). In submission to CHI'10. 2010, ACM.