

Crowdsourcing for Search Evaluation and Social-Algorithmic Search

Matthew Lease
School of Information
University of Texas at Austin
ml@ischool.utexas.edu

Omar Alonso
Microsoft Corp.
Mountain View, CA 94043
omar.alonso@microsoft.com

ABSTRACT

The first computers were people. Today, Internet-based access to 24/7 online human crowds has led to a renaissance of research in *human computation* and the advent of *crowdsourcing*. These new opportunities have brought a disruptive shift to research and practice for how we build intelligent systems today. Not only can labeled data for training and evaluation be collected faster, cheaper, and easier than ever before, but we now see human computation being integrated into the systems themselves, operating in concert with automation. This tutorial introduces opportunities and challenges of human computation and crowdsourcing, particularly for search evaluation and developing hybrid search solutions that integrate human computation with traditional forms of automated search. We review methodology and findings of recent research and survey current generation crowdsourcing platforms now available, analyzing methods, potential, and limitations across platforms.

Categories and Subject Descriptors

H.3.4 [Information Storage and Retrieval]: Systems and software — performance evaluation

Keywords

Crowdsourcing, human computation.

1. Introduction

Human capabilities continue to exceed state-of-the-art automation in many areas, such as computer vision and language understanding. While automation will certainly continue to improve, strategic use of human computation in tandem with automation enables us to build systems which offer greater capabilities today. Use of human computation offers a wider design space in which researchers and developers can explore tradeoffs between processing time, cost, effort, or accuracy. Whereas the traditional mixed-initiative model of human-computer interaction explores the balance between automation vs. user effort, crowdsourcing generalizes this beyond interactions with an individual user to broader social participation. In addition to human-system interactions being driven by an engaged user, we increasingly see interactions where the crowd is engaged by the system to perform work on its behalf. Specific to Information Retrieval (IR), the rise of online crowds has created new opportunities to leverage social wisdom in order to better address information needs than would be possible using purely automated systems and isolated individuals. Today's vibrant crowdsourcing

industry boasts a myriad of vendors offering a wide range of features and workflow models for accomplishing complex, quality work. Researchers and developers can now more easily and effectively create and evaluate search systems with crowds.

While search evaluation is an essential part of the development and maintenance of information retrieval (IR) systems, current approaches for search evaluation face a variety of challenges. Many Web search engines reportedly use large editorial staffs to judge the relevance of web pages for queries in an evaluation set, but this is expensive and has limited scalability. Academic researchers, without access to such editors, often rely instead on smaller test sets than desired, making it harder to detect statistically significant differences in performance. While behavioral data, such as mined from search engine logs of user activity, is much cheaper than the editorial method, it requires access to a large stream of data, something not always available to a researcher testing an experimental system. These challenges provide an ideal setting for demonstrating both the potential and practical challenges for the crowdsourcing paradigm.

2. Target Audience

This tutorial will introduce the opportunities and challenges of human computation and crowdsourcing, including methodology and findings of the latest research, particularly in IR. We will also survey several crowdsourcing platforms, analyzing crowdsourcing methods, potential, and limitations. The tutorial will provide academic and industrial attendees alike with a solid introduction and great start to begin applying crowdsourcing in the context of their own particular research problems or practical tasks.

3. Tutorial Outline

- Introduce crowdsourcing and human computation
- Survey recent “killer apps” of crowdsourcing for search systems (evaluation, training, and crowd-driven search)
- Provide practical “how to” guidance for effectively using several commercial crowdsourcing platforms
- Summarize key best practices for achieving efficient, inexpensive, and accurate work
- Review current opportunities, untapped potential, and open challenges for IR crowdsourcing, particularly evaluation