

Preface

Both crowdsourcing and human computation consider humans as distributed task-solvers, leveraging human reasoning to solve complex tasks that are easy for individuals but difficult for purely computational approaches (human computation) or for traditional organizational work arrangements (crowdsourcing). Effective realizations of these paradigms typically require participation of a large number of distributed users over the Internet, a careful design of task structures, participation incentives and mechanisms for coordinating and aggregating results of individual participants into collective solutions.

Though rarely explicitly addressed as such, social media and related technologies often provide the enabling methods and technologies for the realization of such models. Examples include crowdsourcing marketplaces (e.g. Amazon Mechanical Turk), crowdsourcing service providers (e.g. Microtask, CrowdFlower) or games with a purpose. While centralized platforms are also at the core of “traditional” approaches to collective intelligence (e.g. Wikipedia), attention is increasingly turning to the possibilities of harnessing existing social platforms (e.g. Facebook, Twitter) that already gather huge numbers of users into webs of social relationships. These new approaches to harnessing distributed social infrastructures on the web can enable the design of novel kinds of collective intelligence processes and applications.

Existing Social Clouds (such as Facebook, Twitter) pose both chances and challenges for new kinds of approaches to crowdsourcing and human computation in particular and to collective intelligence in general. On one hand, the intricate social relationships allow the development of new kinds of task routing mechanisms (e.g. identifying the best or most trusted participants for a specific task). Incentive structures are intrinsically social and tend to reflect community-like phenomena (e.g. the reputation economy), thus differing strongly from single-user approaches in classical crowdsourcing. This is already leading to early experiments such as expert-based crowdsourcing or solutions for task-injection across distributed social platforms. On the other hand, the design of such socially distributed computing structures relates the fields of crowdsourcing and human computation to the lessons from a broader class of collective intelligence platforms and applications.

The need to interrelate these fields is reflected in questions such as:

- How can we design effective incentive systems for large-scale participation of human users in structured collective intelligence systems?
- How do we design tasks at different levels of complexity that can be solved reliably through a composition of individual contributions?
- How can we use intricate webs of social relationships of existing social platforms for new models of coordination in distributed task-solving?
- How can distributed social media enable the design of new classes of crowdsourcing applications (e.g. using social network analysis for new ways of task-routing)?
- How can the comparison of lessons from distributed problem-solving in human computation and community-based approaches lead to novel classes of collective intelligence applications?

By explicitly discussing the application of social media to crowdsourcing and human computation as found in practice (i.e. in terms of specific problems and use cases), the exchange of experiences between research communities which are commonly disparate but actually working on related problems can be facilitated. This can result in novel contributions and the identification of best practices in the research areas involved.

Building on the successful workshop SoHuman 2012, this workshop considered crowdsourcing and human computation in the broader context: as specific instantiations of collective intelligence and social computing on the web. The goal was to identify how experience gained from the design of crowdsourcing applications can inform the development of new approaches to collective intelligence, and vice versa: what lessons from the broader domain of collective intelligence can inform the design of new kinds of systems for crowdsourcing and human computation?

Bringing together researchers and practitioners from different disciplines, some of the challenges and opportunities of applying social media to designing novel applications of collective intelligence were explored with a special focus on crowdsourcing and human computation. By emphasizing

practical challenges and experiences with real-world applications, the discussion on best practices in this quickly growing field was stimulated.

This volume contains revised versions of the peer-reviewed papers that were presented at the workshop.

In his keynote, *Bernstein* talks about crowd-powered systems, presenting “Soylent”, a word processor that provides crowd-supported editing tools like shortening or proofreading a paper, and “Adrenaline”, a real-time crowdsourcing application to identify meaningful poses in photographs (both applications interface “Mechanical Turk”). He illustrates how these applications address two of the main crowdsourcing challenges, speed and quality, and proposes a new crowd work design pattern to overcome the latter: “Find-Fix-Verify”.

Hassan et al. discuss how expertise assessment can effect the quality of task routing in human computation. While many crowdsourcing platforms, e.g. “Mechanical Turk”, use pull routing for the distribution of tasks to workers, i.e. workers select tasks they find suitable themselves, the authors propose a push routing approach based on previously assessed worker profiles. Two types of worker profiles are considered, i.e. knowledge profiles based on self-assessment of knowledge about different domains using a likert scale, and performance profiles based on task-assessment as well as a combined approach. For their study, the authors focused on the domains of movie genres and actors and recruited crowd workers via an open international call.

Bry et al. present a crowdsourcing approach for geographical and social mapping of Italian dialects. Making use of gamification and financial market derivative principles, two different games were developed to identify widely and less known Italian dialect/vernacular phrases and to map them to

specific geographical regions or social groups. Users are mostly reached via the authors’ own blog, language community blogs and articles, but the goal is to also target the general public of Italian speaking people.

Fedorov et al. discuss their results from exploiting user generated content for mountain peak detection as a kind of passive human computation approach. They have developed an automatic mountain peak matching algorithm to identify mountain panoramas in geo-tagged user-generated photographs from “Flickr” and “Panoramio”. Their method enables the retrieval of collections of mountain appearances in different time instants and the building of environmental models or new kinds of augmented reality applications.

Dionisio et al. present their approach of both general-purpose and expert-based crowdsourcing in order to build social graphs from historic media collections. Facing the challenge of identifying persons in group photos, the authors combine automatic face detection and recognition with general purpose crowdsourcing via the Mikrotask platform as a validation mechanism of the top 10 automatic face match results. To make better use of domain knowledge in this context, they present a proof of concept to address expert crowds, e.g. historians, journalists and politicians, exploring both implicit crowdsourcing with a shared annotation tool and explicit crowdsourcing using existing social media channels to send out research inquiries.

The workshop was co-located with the ACM Web Science 2013 conference which took place in Paris from May 2nd to 4th. We thank our program committee members, our keynote speaker Michael Bernstein and all authors for their contributions. We also thank the organizers of ACM Web Science 2013 for providing us with the opportunity to organize this event.

August 2013 *Jasminko Novak, Piero Fraternali,
Petros Daras, Otto Chrons,
Alejandro Jaimes, Mark Klein*