

Towards Human-Enhanced Data Management Systems

Alessandro Bozzon (Presenter), Jan Hidders
Web Information Systems group, Delft University of Technology
Mekelweg 4, 2260GA, Delft
The Netherlands

As the amount and variety of data increases, the ability of computers to extract knowledge from the raw data is limited by their lack of cognitive abilities. Machines cannot autonomously elicit meaning, or generate knowledge. In data management systems, human intelligence is employed early (in system design) or too late (in analysing the systems output as a whole): wouldn't it be nice if knowledge-intensive applications could systematically and reliably exploit the intelligence of thousands of brains to perform the small-scale parts of the knowledge extraction problems precluded to machines, and then combine them at large-scale with the effectiveness of computers? To answer these questions, we need a new class of data management systems propelled both by small-scale human intelligence and large-scale power and scalability of computers.

The contribution of this talk is an outlook towards a new generation of human-enhanced systems where crowdsourcing, social networking [1], expertise finding [4], and the organisation (in workflows) of automatic and human activities [3] are blended with the purpose of increasing effectiveness of data management applications. Humans, selected according to their knowledge and skills from crowdsourcing platforms, social networks, or special interest groups, can bring the domain-specific information and the mental agility required to quickly and precisely elicit meaning from raw data. On the other hand, machines can manage and organise at large-scale the outcomes of the small-scale cognitive activities of humans, using the Web as a communication and interaction means, thus guaranteeing speed while reinforcing their artificial intelligence. We report on recent results that show how 1) human computation platforms (such as Amazon Mechanical Turk or Crowdflower) and social platforms (such as Facebook or Twitter), can be used to bridge the cognitive and thinking gap that prevents data management systems from effectively supporting knowledge-intensive and data-intensive applications; and 2) how controlling the quality of human contribution can lead to increased performance and interoperability [2].

References

- [1] Alessandro Bozzon, Marco Brambilla, and Stefano Ceri. Answering search queries with crowdsearcher. In *WWW 2012*, pages 1009–1018, New York, NY, USA, 2012. ACM.
- [2] Alessandro Bozzon, Marco Brambilla, Stefano Ceri, and Andrea Mauri. Reactive crowdsourcing. In *WWW 2013*, pages 153–164, 2013.
- [3] Alessandro Bozzon, Marco Brambilla, Stefano Ceri, Andrea Mauri, and Volonterio Riccardo. Shepherd-ing the crowd. In *Submitted for evaluation to WWW2014*.
- [4] Alessandro Bozzon, Marco Brambilla, Stefano Ceri, Matteo Silvestri, and Giuliano Vesci. Choosing the right crowd: expert finding in social networks. In *16th International Conference on Extending Database Technology*, EDBT '13, pages 637–648, New York, NY, USA, 2013. ACM.