The Challenge of Quality in Social Computation

MILAN MARKOVIC, University of Aberdeen,UK
PETER EDWARDS, University of Aberdeen,UK

CCS Concepts: • **Information systems → Computing platforms**;

Additional Key Words and Phrases: Quality, Social Computation, Provenance, Social Machines, Crowdsourcing

ACM Reference format:
Milan Markovic and Peter Edwards. 2017. The Challenge of Quality in Social Computation. 1, 1, Article 1 (January 2017), 3 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Interactive web technologies now enable a host of so-called *social computations*, which can address challenges that are beyond the capabilities of machines alone. Notable examples of such social computation systems include Galaxy Zoo[1], BeeWatch[2] and Ushahidi[3] - operating in fields as diverse as classification of newly discovered galaxies, monitoring of bee populations, and disaster management. A system for earthquake prediction using social media [5], illustrates how such computations can also emerge on social networking platforms. Social computations can be modeled as a complex collection of structured activities (i.e. workflows) that represent a blend of human and machine tasks, with associated objectives and reward mechanisms. In our previous work [4] [3] we argued that recording provenance of social computation workflows would enhance decision-making support for all associated stakeholders; these include initiators, participants, and beneficiaries of such computations. In the next section, we will briefly introduce the key characteristics of complex social computation systems, before discussing why quality assessments in such a context are challenging.

2 THE CHALLENGE

From a data quality perspective, social computations cannot be treated the same as traditional machine-only computations, as the quality of individual human outputs can vary, even if the same task is repeatedly performed by the same person (due to changing skills, motivation, and opinions). For example, a worker's performance in spotting grammatical errors can improve over time; however it can also decline if the task incentive (e.g., monetary reward) is not sufficient to maintain the worker's attention and motivation. Due to the involvement of hidden human cognitive processes, reasoning about the quality of individual human contributions based on their characteristics and actions during a computation poses a significant challenge. However, provenance

[1]http://www.galaxyzoo.org/
[2]http://homepages.abdn.ac.uk/wpn003/beewatch/
[3]https://www.ushahidi.com/

The research described here was supported by the award made by the RCUK Digital Economy programme to the dot.rural Digital Economy Hub; award reference: EP/G066051/1.
Author's addresses: M. Markovic (milan.markovic@abdn.ac.uk) and P. Edwards (p.edwards@abdn.ac.uk), Computing Science, University of Aberdeen, United Kingdom, AB24 5UA.
2017. XXXX-XXXX/2017/1-ART1 $15.00
DOI: 10.1145/nnnnnnn.nnnnnnn

information detailing the design and execution of a social computation can usefully aid quality assessments, and is easier to document. For example, Huynh et al. [2013] describe how records of an entity's influence captured using the World Wide Web Consortium recommended data model for describing provenance on the Web (PROV)[4] can be used to construct a predictive model indicating the likely quality of data generated by a social computation. Unfortunately, PROV is unable to describe some of the crucial characteristics of social computation workflows. To illustrate, consider the discussion in Bernstein et al. [2010] in which they observed that the overall structure of a social computation, such as the order and the nature of individual tasks, and the type of aggregation mechanisms used (i.e. a workflow plan) can affect the quality of the generated output. While PROV can document the workflow execution, it cannot describe details of a social computation workflow plan at the level of individual tasks and their associated characteristics.

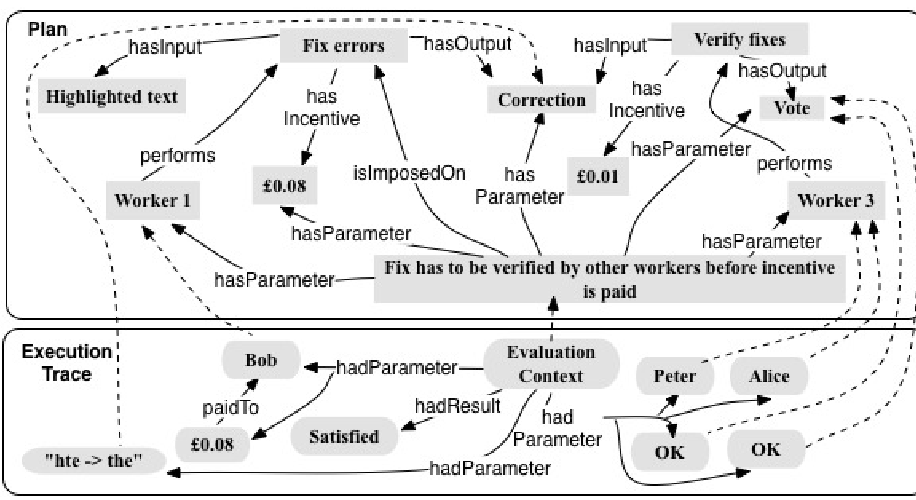

Fig. 1. Part of the *CrowdProof* workflow used in Soylent [1] to fix grammatical errors in a piece of text enhanced with a single reward condition.

To illustrate the importance of capturing detailed descriptions of plans in the context of social computation, consider an example based on the *CrowdProof* component of the Soylent word processing tool [1]. In *CrowdProof*, workers perform three types of activity, namely finding text patches that include grammatical errors (*Find step*), proposing fixes (*Fix step*), and verifying proposed corrections (*Verify step*). In the first two of these, workers are required to create new content (e.g., a fix for a grammatical error), while the verify activity requires a different set of workers to decide (as a group), which of the proposed fixes should be accepted as a solution to the overall task. Each of these steps can be associated with various constraints influencing aspects of the workflow execution, such as worker eligibility, number of workers required to perform individual tasks, and conditions associated with incentive payments.

Figure 1 depicts two plan steps representing human tasks (i.e. *Fix errors* and *Verify fixes*), their associated incentives, and a reward condition associated with the *Fix step*. This reward condition restricts the payment of incentives, so that they are only awarded to workers who had their solutions verified by two others in the *Verify step*. The graph representation also includes a sample execution

[4]https://www.w3.org/TR/prov-dm/

trace describing the enactment of the workflow plan. Figure 1 highlights many of the important characteristics of this social computation that could be utilised by quality assessment metrics. For example, such methods might consider the presence of a particular condition (e.g., restrictions on a worker's country of residence), or the approach used to aggregate and validate individual solutions, or deviations from the original plan recorded in the execution trace (e.g., some parts of the workflow did not execute due to inadequate incentives, overly restrictive conditions, etc).

3 RESEARCH DIRECTIONS

We argue that the following represents a pressing challenge for information quality researchers: *How should the quality of social computation outputs be assessed based on the various structural and operational characteristics of such a computation?*

To address this challenge, further extensions to current provenance models may be needed together with new forms of quality metric. To begin addressing the former, we designed a PROV extension called SC-PROV [3], which represents a first step towards a social computation provenance model that includes a record of planned tasks, incentives, and associated constraints (e.g., required worker qualifications, geographical location, reward constraints) in addition to agents, activities and entities that are required to describe the execution of a social computation. Further research should also focus on capturing and analysing real provenance records describing the *expected* as well as *actual* behaviour of a social computation to discover relevant provenance features (e.g., specific parameters of task constraints) and interdependencies among these features, which could inform future quality assessment methods.

Social computation systems raise unique challenges in terms of quality assessment of their outputs - due to their use of interconnected human participants on a global scale. This necessitates not only an inquiry into the suitability of traditional quality models designed for machine-only systems, but also creates exciting opportunities for research into new forms of quality assessment that will require input from researchers in fields such as computer science, economics, psychology, management, sociology, and others.

REFERENCES

[1] M.S. Bernstein, G. Little, R.C. Miller, B. Hartmann, M.S. Ackerman, D.R. Karger, D. Crowell, and K. Panovich. 2010. Soylent: a word processor with a crowd inside. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. ACM, 313–322.

[2] T. D. Huynh, M. Ebden, M. Venanzi, S. D. Ramchurn, S. Roberts, and L. Moreau. 2013. Interpretation of crowdsourced activities using provenance network analysis. In *First AAAI Conference on Human Computation and Crowdsourcing*.

[3] M. Markovic. 2016. *Utilising Provenance to Enhance Social Computation*. Ph.D. Dissertation. University of Aberdeen.

[4] M. Markovic, P. Edwards, and D. Corsar. 2013. Utilising Provenance to Enhance Social Computation. In *The Semantic Web–ISWC 2013*. Springer, 440–447.

[5] T. Sakaki, M. Okazaki, and Y. Matsuo. 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors.. In *Proceedings of the 19th International Conference on World Wide Web*. 851âĂŞ860.