CEST
Centro de Estudos Sociedade e Tecnologia
USP
Universidade de São Paulo
Bulletin - Volume 1, Number 10, April/2017

Using *Crowdsourcing* for the Benefit of Society

Eduardo Bertassi

In 2009, the House of Commons of the United Kingdom experienced one of the biggest political scandals in its history when *The Daily Telegraph* had access to more than two million documents containing information evincing that British Members of Parliament (MPs) were using public money to pay back personal living expenses. The newspaper started reporting the irregularities committed; however, the reports were gradually made due to the huge amount of information that had to be checked out before made public.

The scandal grew to bigger proportions when *The Guardian* had access to the same documents and posted them in a website called "*Investigate Your MP's Expenses*", which allowed any logged-in user to access, analyze, and indicate any suspected information that should draw the journalists' further attention.

The term *crowdsourcing* had the literal meaning of *outsourcing* the execution of a certain activity to a *crowd* scattered in the internet.

In the first 80 hours in which the website was available, 170,000 documents were reviewed. In the whole, more than 20,000 people freely helped journalists of *The Guardian* to find and report MP-related irregularities, leading to the resignation of the Speaker of the House of Commons, Michael Martin.

This is just one of the several examples that may be given about the use of the *crowdsourcing* model in benefit of society.

When it first appeared, the term *crowdsourcing* had the literal meaning of *outsourcing* the execution of a certain activity to a *crowd* scattered in the Internet. Although the concept of collaborative work is not new, the credit for creating the term belongs to journalist Jeff Howe through a report elaborated for *Wired* magazine in 2006.

The intensive use of smartphones, mobile applications, and social network along the years made people become *prosumers*: individuals who are both *producers* and *consumers* of contents. Consequently, new *crowdsourcing*-based business opportunities and *software* have emerged, such as *Foursquare*, software used for recording users' preferences that allowed new findings about collective human behavior; *reCaptcha*, software that, by means of micro tasks, helps with the process of digitalization of several books; and *Kickstarter*, website that facilitates raising funds to carry out new projects.

In 2015, researchers from the Faculty of Science & Technology at the University of Bournemouth, United Kingdom, published a work that demonstrated the possibility of categorizing problems or tasks that could be resolved by using *crowdsourcing* in five types:

- **Gathering Opinions** - the objectives of the tasks were only to gather opinions (there were no right or wrong answers);
- **Simple Problems** - the objectives of the tasks were easily achieved (such as collecting photos or counting), since they did not require specific knowledge for their execution.
- **Complex Problems** - the tasks were more complex, since they required specific knowledge of further cognitive capacities for their execution (such as writing a document together, develop software in team work, among others).
- **Competitions** - the tasks could be easy or difficult to be carried out, but only one participant (or one group) would be rewarded should they be successful.

- **Fundraising** - the tasks consisted in raising funds (monetary or not) non-compulsorily to help a charity institution or to carry out a project.

The first step to create some problem-solving *software* using a *crowdsourcing*-based approach is to identify which of the five categories the problem belongs to. After that, for each identified category one must answer four basic questions related to the possible interaction between people and the *software*:

1. **Recruiting** - How to recruit and retain users?
2. **Task Design** - How to define and design tasks so that people can be suitable for executing them without losing motivation?
3. **Combination of Efforts** - How people's contribution can be combined so that the results can be maximized?
4. **Management of Activities** - How the execution of tasks can be managed and judged to prevent the final result from being incorrect?

The Guardian's *software* developers had to be creative when transforming the Complex Problem into a Simple Problem with tasks that could be executed by several unknown users scattered in the Internet and with different instruction levels. When seeking answers to the four basic questions presented, they identified that:

1. **Recruiting** - Since it was a polemic, recent issue, present in people's mind, they developed a funny *software* that would stand out the one who had contributed more, since they could not pay or reward the participants directly;
2. **Task Design** - The tasks would have to be as simple as possible. People should only indicate whether the information: a) was interesting (having relevant expense data); b) was interesting, but familiar (having relevant expense data, but within a normal standard; c) was not interesting; or, d) deserved further investigation;
3. **Combination of Efforts** - People should manage to analyze any documents independent of their ordination. The resulting collective effort would allow finding each MP's individual information;
4. **Management of Activities** - The information that deserved further investigation should be checked out by *The Guardian*'s team to avoid mistakes that could compromise the newspaper credibility.

The Guardian's development team's agility, together with the social upheaval of the moment (and a bit of luck), allowed the website to be successful. From the *software* engineering viewpoint, it is desirable that this success should be replicated through areas, such as, environmental sustainability, urban planning, identification of catastrophes, and even space exploration minimizing risks, costs, and implementation time without compromising the final quality of the solution.

It is known that, usually, the questions presented are hard to be responded to in time of project due to the complex, dynamic, and unpredictable behavior of thousands of users scattered in the network, able or not to interact with the *crowdsourcing software*. Cases such as that of *The Guardian* are inspiring, but there are also familiar cases in which the *crowdsourcing* model was misused resulting in a waste of workforce, time, and financial resources of the companies.

It is expected that the uncertainties and difficulties when developing some *software* for *crowdsourcing* should be seen as challenges and opportunities for researching, for example: people's intrinsic and extrinsic motivations when executing tasks through the internet; more intuitive interface construction technique to execute more complex tasks; automation of mechanisms to create and control tasks; better forms of rewarding the most committed users; among others.

Improving the development of *crowdsourcing-based software* can be one of the paths that will allow technology - allied to people's will - to help build a fairer and transparent society. Dishonest people should be careful.

Eduardo Bertassi *is a master's degree student in computing engineering at Escola Politécnica da Universidade de São Paulo, and researcher at CEST-USP.*

Coordinator: Edison Spina

This article is a result from the author's ascertainment and analysis, without compulsorily reflecting CEST's opinion.