



## How to Use Crowdsourcing Effectively: Guidelines and Examples

**Elena Simperl**

University of Southampton, United Kingdom  
e.simperl@soton.ac.uk

### Abstract

Crowdsourcing has become the *mot du jour* when it comes to resolving any types of problems, online or offline, that require sustained human involvement. We see it applied in order to motivate employees to engage with less rewarding daily routines, to attract the best possible ideas and approaches to boost innovation, or to complete data processing tasks that computing technology has yet to master quickly and accurately.

In this paper we look at its various forms and flavours, from gamification to human computation and grand challenges, and discuss how it could be used to turn conventional content management applications into social machines in which tasks are performed as optimal combinations of human and computational intelligence. We introduce a framework for the analysis of the most important building blocks of such systems, as well as design and participation best practices that should guide their development.

**Key Words:** crowdsourcing; content management

### 1. What is Crowdsourcing?

‘Crowdsourcing’ refers to a novel approach to distributed problem solving, in which tasks traditionally assigned to the employees of an organization or to a designated group or community of interest are outsourced to a loosely

defined 'crowd' of people through an open call (Howe, 2006). This can take various forms, from peer production (Haythornthwaite, 2009), in which work is undertaken collaboratively, to public competitions, in which only the best contributions will be recognized, to participatory sensing, which uses mobile devices and sensor networks to collect vast amounts of information (Burke et al., 2006). Less than a decade after the introduction of the term in 2006, we see crowdsourcing applied to virtually any domain, from using gamification to drive employees' motivation, to challenges and prizes rewarding ideas for product development and innovation, to paid microtasks as a new form to complete routine content work such as simple text translation, data entry, or updates of database records (Dawson & Bynghall, 2012).

Some could rightfully argue that the concept at its core – achieving a goal via contributions from many individuals – is not that new; in fact some of the most successful exemplars of the 'wisdom of the crowds' (Surowiecki, 2005) are dated pre 2006 (e.g., Wikipedia), or even, for those who remember, pre Web and information technologies. 'Online crowdsourcing', however, is still quite a recent phenomenon and the efficiencies it brings in mobilizing many people in a relatively short period of time make it useful in many ways. In the following we will elaborate on the challenges and opportunities it creates.

## **2. Online Crowdsourcing: Challenges and Opportunities**

As the name suggests, the rise of the Web, smart phones, and affordable wireless sensors meant that organizations interested in crowdsourcing could easily reach out to a global pool of resources, skills, and creativity, readily available at almost any time of the day at the click of a button. In other words, when we talk about crowdsourcing, we typically think about scenarios involving groups that are orders of magnitude larger than in the classical teamwork scenarios that have been subject to organizational management and collective intelligence studies (Malone, Laubacher, & Dellarocas, 2010). The scale of the exercise has implications for the ways in which individual contributions are planned, assigned, coordinated, and appraised; a 'typical' crowdsourcing project would involve work that can be broken down into many smaller chunks that can be completed in the same time by different parties; and evaluated to a large extent automatically. This additional overhead could potentially outweigh the expected benefits of the exercise; for

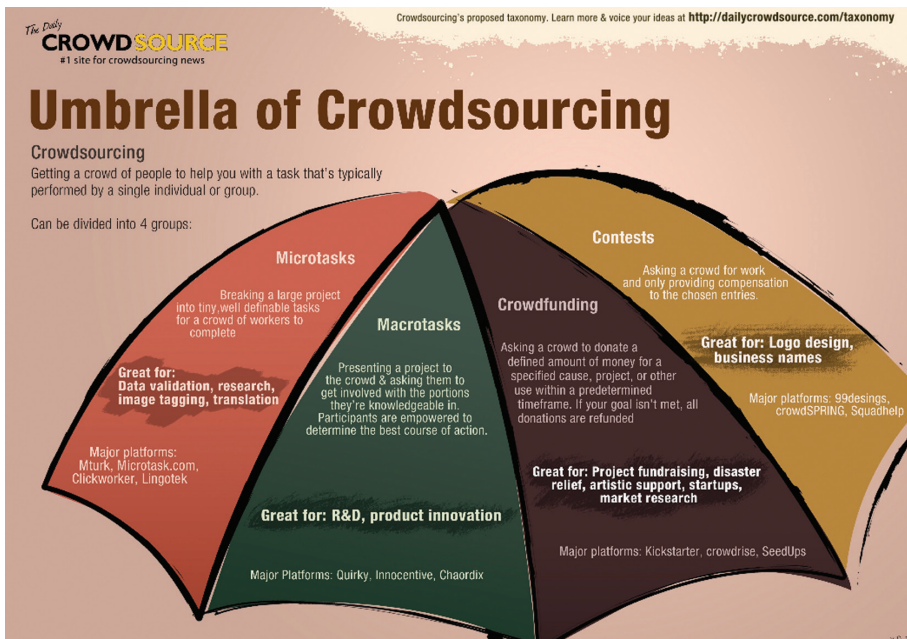
instance, when Google launched a project asking for ideas on how to make the world a better place, it took them around three years and 3,000 employees to review the 150,000 submissions and identify 16 projects they eventually pursued.<sup>1</sup> The Netflix challenge is another great example. While the contest attracted a fair amount of attention in the media and received several valuable contributions, the actual results were never used because the engineering effort required to integrate the winning algorithm into the platform did not match Netflix' new business model and shifts in customer behaviour. While the initiative most likely added to Netflix' image as a technology pioneer, discrepancies in the definition of the crowdsourcing task and evolving technical and economic conditions meant that the creative investment of the challenge participants was mostly lost.<sup>2</sup>

A second critical development in online crowdsourcing was the new culture of openness, which gradually changed the way government and the private sector think about engagement and of the role that citizens and customers – individuals or larger crowds alike – could play in improving innovation potential and public image. As a central element of crowdsourcing, the open call has both upsides and downsides. The upsides are a function of the sheer scale of the audience targeted by the call: organizations have access to vast pools of external spare resources, sharing risks and rewards with others for mutual benefit. The downsides become apparent when touching upon IP areas that are too critical to disclose, or because of the lack of insight and institutional practice for assigning tasks to the most suitable crowd members and incentivizing their behaviour. Far away from the familiarity of established social structures, organizations are confronted with a great potential, but also with an opaque mass of loosely committed contributors; little is known about who they are, what they are good at, and what motivates them in short, medium, or long term. The research community and crowdsourcing service providers have put considerable effort into designing methods to identify and discourage spam (e.g., Oleson et al., 2011) and retain contributors' engagement (e.g., Cuel et al., 2011).

Finally, platforms such as InnoCentive, oDesk, Ushahidi, Kickstarter, or CrowdFlower<sup>3</sup> have greatly simplified the execution of crowdsourcing projects, bringing together 'requesters' (the person or institution seeking help from the crowd) and 'workers' (individuals or teams taking on tasks advertised via an open call). Putting aside the principled differences between forms of crowdsourcing (see also Figure 1), which make them from the offset

amenable only to specific types of problems, the success of any crowdsourcing endeavour will crucially depend on the ability of the requester to benefit from crowd inputs and encourage participation, while taking into account frame conditions such as time, budget, and quality. For instance, one tends to distinguish between macrotask and microtask crowdsourcing scenarios to refer to the granularity of tasks that are accomplished by individual contributors. Microtask crowdsourcing is used when the task is highly parallelizable and can be divided into smaller pieces down to a ‘micro’ level, which takes only seconds to minutes to complete. The model is very similar to the MapReduce programming paradigm, according to which a task is executed via a parallel, distributed algorithm. Macrotasks, in turn, refer to those cases when the Map phase of the algorithm is not straightforward to define. This is typically the case for creative tasks or other complex tasks whose resolution requires a great share of contextual information or dependencies to intermediary results. Another important feature to distinguish forms of crowdsourcing is the incentive mechanisms that are put in place. Both macro- and

Figure 1: Main forms of crowdsourcing according to The Daily CrowdSource.<sup>4</sup>



microtask projects could rely on any combination of the money, passion, glory triangle (Malone et al., 2010). Requesters could organize a contest and reward the best contributors, appeal to people's intrinsic motives to participate freely, or pay each of them based on their efforts. The challenge then is in the mismatch between these factors and the level of support offered by crowdsourcing platforms, which is more often than not rudimentary. Most of these platforms merely provide access to their crowd of registered contributors, leaving the requesters alone with the design of the actual project, the definition of rewards, and the assessment of the crowd outputs. Later in Section 5 we will introduce a principled approach to guide this process, informed by existing case studies, related literature on incentives and motivation, and our own experience. We propose a four-dimensional analytical framework, which assists a requester in her effort to identify promising ways to reach out to a crowd and benefit from the outcomes it produces. To illustrate the use of the framework, we will refer to two examples familiar to the research libraries community: one related to citizen science, a novel approach to pursue scientific inquiry using the knowledge and skills of non-expert volunteers (Wiggins & Crowston, 2011); and a second one around scholarly publishing, in which we explore the use of crowdsourcing to enrich existing metadata about research papers with links to the datasets they cite (Drăgan et al., 2015). However, before we go on to introduce the framework, we will first add more context to the ways it could benefit researchers and library practitioners by positioning crowdsourcing against related terminology and then discussing a set of existing crowdsourcing projects that we deem relevant to the field.

### **3. Related Areas**

Since the invention of the Turing machine we have experienced a paradigm shift in the usage of computers, which slowly advanced from purely calculative devices to facilitators of a wide range of human interactions. The emergence of 'Computer-supported Collaboration Work (CSCW)' (Grudin, 1994) is representative for the early days of this trend. CSCW addresses "*how collaborative activities and their coordination can be supported by means of computer systems*" (Grudin, 1994), whereas the initial concept evolved towards the more broader field of 'Computer-supported collaboration' (CSC), which encompasses work-related aspects.

The Web as a global platform for information access and sharing marked a second essential milestone; in particular, through the principles and technologies promoted under the label Web 2.0 and the proliferation of smart mobile devices. These developments led to an amazing growth in terms of the amounts of content available online and mass participation. They are responsible for hundreds of millions of users all over the globe creating high-quality encyclopaedias, publishing Terabytes of multimedia content, contributing to world-class software, and lively taking part in defining the agenda of many aspects of our society. This progression towards 'prosumerism' found more and more adopters in the public and private sector as well, as governments and enterprises not only became active in open initiatives, but sought the knowledge and advice of their customers and employees in taking decisions related to organizational management, product development, services offers, and policies. In this context, a number of terms are used to refer to the ways people interact with each other and with applications: 'wisdom of the crowds' (Tapscott & Williams, 2008), 'collective intelligence' (Lévy & Bonomo, 1999), 'open innovation' (Chesbrough, 2003), 'human computation' (Von Ahn, 2009), 'social computing' (Wang, Carley, Zeng, & Mao, 2007), and 'social machines' (Hendler & Berners-Lee, 2010).

Wisdom of the crowds (Surowiecki, 2005) refers to a principle for decision making that takes into account the input of a group of people rather than just individuals; the use of specific technologies, most notably Web 2.0 and mass collaboration tools, has made it possible for such processes to be carried out at scales hardly conceivable in the past, and to involve highly diverse and geographically distributed participants. A similar concept, though broader scoped, is collective intelligence, defined in (Malone, Laubacher, & Dellarocas, 2009) as "*groups of individuals doing things collectively that seem intelligent*"; the field is concerned with all forms of collective behaviour, including animal and artificial intelligence. The crowd element in crowdsourcing has commonalities with both these areas. Compared to the wisdom of the crowds, it studies a slightly more focused class of scenarios. We are talking about a requester aiming to achieve a certain goal via contributions submitted in response to an open call. Sometimes, individual crowd inputs will be coalesced into a final output; other times, only a selection of these inputs will be deemed as useful and recognized accordingly. The collective component in crowdsourcing can be more or less explicit. The emphasis is on human participants, supported by technology.

One of the direct consequences of the popularity of the wisdom of the crowds notion was a stronger investment worldwide in open innovation. Open innovation could in fact be seen as a manifestation of the wisdom of the crowds in business environments, or, in the words of the authors, as a *“paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology”* (Chesbrough, 2003). At a more general level, the approach could be applied to any domain, from science to public policy, and crowdsourcing is typically associated with this large collection of scenarios. Just as in open innovation, the crowd responds to a call in which contributions are sought to achieve a specific goal.

When these goals are heavily motivated by technology we speak about human computation (Quinn & Bederson, 2011). More specifically, human skills are applied to tackle technical tasks that computers still find challenging, for example, summarizing or paraphrasing text or recognizing things in images. Unlike open innovation, the focus is on so-called ‘microtasks’, which refer to simple works that require basic language understanding and cognitive processing capabilities in the range of seconds to minutes to complete. This sort of tasks are an important part of today’s crowdsourcing landscape, in particular via platforms such as Amazon’s Mechanical Turk<sup>5</sup> or CrowdFlower, which offer small financial rewards to an anonymous crowd engaged with microtasks posted by various requesters. Human computation is also at the core of a research area called ‘games with a purpose’ (GWAP) (Von Ahn, 2009), which build a game-like environment including points, badges, leaderboards and other common game elements to encourage people to complete microtasks.

Social machines refer to online socio-technical systems composed of crowd and algorithmic components (Smart, Simperl, & Shadbolt, 2014). Compared to crowdsourcing, there is less of a focus on an open call inviting contributions towards a specific goal. Many content sharing platforms and social networks are great examples of social machines, though their outsourcing element is less pronounced. A more important difference between the two is hidden in the automation part of social machines, which are about principled and useful ways to combine human and computational intelligence and not just about accomplishing goals with the help of an open crowd. A similar line of reasoning could be followed to point out the overlaps between social computing and crowdsourcing. The latter is an area of computer science which



refers systems that support “*the gathering, representation, processing, use, and dissemination of information that is distributed across social collectivities such as teams, communities, organizations, and markets*” (Parameswaran & Whinston, 2007). As such, compared to the general concept of ‘Computer-supported collaboration’, social computing puts a greater emphasis on the information management capabilities of groups and communities, and less on the way these capabilities emerge as a joint effort. Crowdsourcing benefits from social computing technology to build useful tools that support crowdsourcing projects, from marketplaces where requesters and workers meet to methods to encourage collaborations and ideas exchange, coordination, engagement and results validation.

#### 4. Crowdsourcing and Research Libraries

There are several classes of crowdsourcing applications pertinent to research libraries that have already proven successful. First, there are platforms such as figshare<sup>6</sup>, ResearchGate<sup>7</sup>, Mendeley<sup>8</sup>, and Taverna<sup>9</sup>, which offer a forum for scientists to manage, share, and comment upon research outputs such as publications, background literature, datasets, and experimental workflows. From a purely crowdsourcing standpoint, common to these initiatives is a strong online community element. The basic assumption is that the main incentive for people to register to and use the platform is the services it provides, which in the long run would lead to more visibility among like-minded peers and a higher impact of one’s research. From a content management point of view, participants are asked to contribute information about their own publications and related research artefacts, manage shared lists of references, exchange ideas and commentary, and network. A second class of applications focuses less on scholarly publishing, but aims at building communities of interest around a particular scientific topic. There are various examples in this space, from cultural heritage projects such as History Pin<sup>10</sup> to citizen science platforms such as Zooniverse<sup>11</sup>, hosting tens of individual projects in natural sciences and the humanities. In this context we also find open innovation-centric R&D platforms such as InnoCentive<sup>12</sup> and Kaggle<sup>13</sup>, which outsources hard scientific problems to a global community of experts and online communities of practice such as PolyMath,<sup>14</sup> which encourages scientists to work together and solve problems of general interest. Finally, there are initiatives and tools that are not specific to digital libraries, but could nevertheless prove



useful in supporting research librarians. This third area is very broad, and encompasses anything from crowdsourced encyclopaedias to projects like Distributed Proofreaders<sup>15</sup>, which uses human computation to improve OCR text detection, or Duolingo<sup>16</sup>, which advances machine translation by gamifying language learning.

From the diversity of these systems we learn two things: first, that crowdsourcing can be successful in many areas; second, that there are many ways in which one could apply a crowdsourcing approach to achieve a specific goal. In the following we will introduce a framework which structures the crowdsourcing design process and helps potential requesters in understanding the differences between the different forms of crowdsourcing and the implications of choosing one or another.

## **5. A Guide to Crowdsourcing**

### **5.1. Overview**

In any crowdsourcing endeavour one speaks about a requester, which is the person or organization that issues the call to outsource, and a general goal to be achieved through the crowdsourcing activity. This goal will translate into specific tasks that members of a crowd are invited to take on. The outcomes they achieve are then assessed and potentially assembled into a final result that the requester can use. Participation of the crowd relies on the motivations of the contributors as well as on incentives mechanisms engineered by the requester.

Based on this simplified view of a crowdsourcing process, we distinguish between the following four design dimensions in our framework:

- 1. What to crowdsource.** A mapping of the high-level goal to specific tasks to be completed by the crowd on specific platforms.
- 2. Who is the crowd.** Crowdsourcing is essentially driven by an open call, which implies that the requester knows little to nothing about potential contributors. However, practically the crowd that can be reached via the call and should respond to it is determined by the crowdsourcing platforms and advertising channels the requester uses or by knowledge and skills pre-requisites.

3. **How to crowdsource.** Depending on what is crowdsourced and the foreseen participants, the requester has to design and execute the tasks and define assessment criteria and put tools in place to consolidate individual contributions into a result that can be used by the requester. There are a number of options to make the process more effective, reducing unintended behaviour (e.g., spam), and exploiting economies of scale, talent, availability, time, and budget.
4. **How to incentivize.** Participation is an essential pre-requisite for any crowdsourcing project. Task, crowd, and platform features have an impact on the number of contributors and their level of performance and engagement. Choosing the right incentive mechanisms and refining them to shape crowd behaviour greatly contributes to the success of a project.

In the following we will elaborate on each of these dimensions and bring in examples relevant to the research libraries community.

## 5.2. What to Crowdsource

While the overall goal of a crowdsourcing exercise might be clear, it is often too general to be translated one to one into an open call targeting a large unknown group of people. The basic assumption is that the reason why a requester considers crowdsourcing in the first place is because it is more effective (in terms of time, costs, or quality of results) than other options – these being, the knowledge, skills, and expertise available in-house, or the use of ICT tools. Open innovation would target the former; human computation would be related to the latter. More specifically, we see it applied to those tasks that require knowledge, cognitive, and social skills that are not easily replicable by computing technology. Typical examples can be found in the creative industries (e.g., ideas for new products, logo designs), in R&D (e.g., improving Netflix' recommendation algorithms), or in the broad range of common daily activities humans are exceptionally good at (e.g., understanding written or spoken language, identifying and classifying objects in pictures, finding and cross-checking facts, assessing subjective qualities of content such as sentiment, aesthetics etc.).

Putting aside the broader debate about the natural or present limits of artificial intelligence, decisions to use crowdsourcing to carry out a specific

tasks are often a matter of resources: finding the right software which would offer a similar functionality, customizing it to one's needs, training it using specific data and analysing the results. The reason why some many requesters prefer human to machine computation is because in many cases at least one of these activities is not effective. GalaxyZoo, the first Zooniverse project, used people to classify galaxies in images because putting together a software alternative would have taken months and substantial investment; building on clever promotion and the natural fascination people have for the universe, they completed the task in a matter of days.

Going a step further, it is important to think about the best way to translate the high-level goal of the project into specific tasks that can be presented to the crowd. This boils down to at least two important questions: what will the crowd see, or how would one create an easy-to-understand description of the task at hand, and what kind of contributions does one expect. We will explain both using an experiment in the area of scientific data citation (Drăgan et al., 2015; Hitchcock et al., 2002). The ultimate aim was to add information about data sets that scholarly publications refer to, as a means to improve research reproducibility and get a feeling of the impact of specific data sets. One could imagine a publisher wanting to offer richer services to its readers through augmented content, or a scientist who shared a data set with the broader community looking for a way to demonstrate impact beyond her publications and the number of citations they achieve. This general goal could be understood in several ways from a crowdsourcing point of view:

1. **Given a publication, tell me which data sets it mentions.** In this case, the requester is aiming to collect a list of data sets from the crowd. A list of publications is enough to kick-start the crowdsourcing process. However, very little is known about which answers would be correct or not due to the open nature of the question. This possibly complicates the use of the results, especially the answer set is large.
2. **Given a publication and a data set, tell me if the data set is relevant to the paper.** This is a much more constrained task, with a closed set of possible answers, one of which would be correct: either the paper mentions the data set or not. This means that the requester would most likely have less trouble assessing the contributions (more about this later), but is also assumes that a list of potentially relevant data sets is available in advance.

3. **Given a publication and a list of data sets, choose the one(s) that are relevant.** This is variation of the first two that accepts multiple answers and requires that the requester knows which data sets could be subject to the task at all.
4. **Here is a data set, tell me a paper that mentions it.** This is the data set-centric version of the first task, with similar characteristics and a possibly even larger answer set.
5. **Here is a data set and a list of publications, tell me which publications refer to the data set.** Again, a multiple-choice variation of the previous task, which would assume that the requester can generate a plausible list of candidates for the crowd to choose from.

Complementary to the distinction between open and closed tasks one has to consider its interface to the crowd. No matter which of the five tasks one would go for, the next step would be to decide how the two relevant types of items (publications and data sets) would be shown to the crowd. Publications are text and as such one could consider anything from displaying title and authors to the abstract, the first page, or even the entire document. Other forms of media are less straightforward to render. In particular for data sets, one could mention the name of the data set or a Web site that describes it. Whatever the human interface is eventually selected, the aim should be to balance the time required to complete the task with the knowledge and context required to do it well. For example, reading a full page of text might give the contributor lots of information about the research and its data sets, but it will take time. If data sets tends to be mentioned in the abstract, then showing only that restricted view of the publication to the crowd might do. Related to this is also the question of identification: we could imagine two versions of a paper with the same title and author (e.g., a technical report and a peer-reviewed article), or even more so, multiple versions of the same data set. The requester needs to be aware of these details and decide whether they are part of the answers he is seeking to obtain from the crowd.

There might also be an issue with the availability of task inputs. The full paper might not be available for free for the contributor to consult, while the title, authors, keywords, and abstract typically are. A selection of data sets or publications to choose from would have to be built in advance, possibly using automatic tools, which have their own limitations. Giving the crowd a set of options makes the results more predictable and as such easier to process, but

it also means that if the initial list was greatly incomplete or just wrong, the results from the crowd will reflect that.

Finally, it might be that the initial goal will be translated into a number of tasks to be carried out by different crowds, possibly as part of a greater ecosystem involving in-house expertise and automatic tools. For example, one could imagine a scenario in which our scientific publisher would first run an information extraction algorithm on a corpus of publications to identify names of data sets, then run some crowdsourcing experiments on Mechanical Turk to filter out false positives and, complementarily, to find out missing data sets, before finally launching a Twitter campaign calling publications authors to double check the results and add versioning information.

### **5.3. Who is the Crowd**

This example is also illustrative of the role of the crowd in a crowdsourcing project. Each specific task will require skills and know-how, some more difficult to find than others. This has to be taken into account when deciding how the task will be presented to the crowd. For example, in GalaxyZoo citizen scientists are only gently introduced to expert terminology – the aim of the Zooniverse designers is to phrase the questions in a way that is understandable without having any background in astronomy in order to appeal to a greater group of people. However, many classes of crowdsourcing projects target specific communities of interest, either due to the nature of the task itself, or the way it will be crowdsourced (see next section). Sometimes the platform dictates the audience. figshare and Mendeley are explicitly geared at researchers. Duolingo is appealing to people who want to learn a foreign language. In games with a purpose the human computation element is hidden underneath a game narrative; they target primarily casual gamers. In other cases, the channels used to promote the open call implicitly introduce a bias in the formation of the crowd. Social networks will reach only people directly or indirectly related to the account making the post. Mailing lists and discussion forums, even when they are open, are frequently read only by a specific group of subscribers; the same applies to Web sites. It is important to understand which types of crowds could be potentially relevant and necessary to achieve the requester's goal. Some of them will be accessible via dedicated channels; they will engage with the task differently and be more or less motivated to reply to the call. Even when they do, the quality of the outcomes

will vary greatly and, as we are talking about non-traditional work environments, the requester will initially have very little knowledge about skills, availability, and willingness to contribute. As we will see in the next sections, there are ways to compensate for this lack of information, either by using a particular platform, by learning to predict performance of crowd members from previous interactions and by aligning incentives and motivation.

#### **5.4. How to crowdsource**

Earlier we discussed several factors that help us understand the main classes of crowdsourcing. One important dimension is the level of granularity of the task.

In a macrotask scenario, the task is outsourced as it is to one or more contributors. The assumption is that the requester will evaluate the submissions and select a small subset of them. This model is followed in open innovation, in challenges, in participatory government, or on virtual labour marketplaces such as oDesk. The requester needs to define the task, evaluation criteria, and incentives. There are a number of options both for assessing and rewarding contributions (Quinn & Bederson, 2011). Assessment could be carried out manually, by the task owner or by a panel of experts, openly or privately. It could involve votes from the community, including peer contributors, or just an automatic algorithm. All this affects the level of engagement and the degree to which the crowd is willing to subscribe to the final decisions. This aspect is important in particular for those types of tasks which are difficult to evaluate with objectivity, from ideas co-creation to product design. Regarding incentives, a great deal of crowdsourcing projects rely on volunteers. Citizen science is just one prominent example, but one could look at any platform for user-generated content or at open-source software projects as well. Challenges like the Netflix one work with sometimes significant financial rewards; the same applies to virtual workforce providers.

In a microtask scenario, the task will be broken down into smaller chunks executed independently. In some cases this might be less straightforward than one thinks. Consider, for instance, a language translation project: a document of hundreds of pages should be translated from one language to another. One option would be to use an online service, which will deliver a complete translation in one go, possibly requesting several quotes from the

crowd to get an idea of the level of quality in advance. The alternative would be to break down the overall exercise into bits of text that can be translated in a matter of minutes by multiple translators, and merge the results into the final document. Using a platform like CrowdFlower one could complete the task in a matter of hours, but additional effort would be needed to compare alternative translations and make sure the integrated document is consistent and reads well. In the same vein, this second step could be divided into subparts of several pages each, which would be proof-read and edited by the crowd. Having several people work on the document at the same time would expedite the process. The important point in this example is the fact that achieving a higher efficiency might require translating the initial task in complex workflows, which require coordination, and additional design effort.

A second dimension distinguishes between explicit and implicit crowdsourcing. In the first case, the requester uses a professional crowdsourcing platform as well as social media and other PR channels to solicit contributions. In the second case, the purpose is less explicit. This applies most prominently for games with a purpose or for participatory sensing. In both examples the crowd does not explicitly join solve tasks; they play a game or collect and share information via their mobile devices or other sensors, and the results are used by the requester for their own purposes. These examples should make clear that the basic approach – achieving an aim or solving a problem with the help of a group of people – is not new. In fact any type of social computing technology, from Google updating their search algorithms based on who is clicking on which links to Amazon recommending products using collaborative filtering has similarities with implicit crowdsourcing. However, what is new in crowdsourcing is the explicit call, including goals, assessment criteria, and rewards, which are not central to social computing. From a crowdsourcing point of view, the fact that collective action is taken implicitly often means a change in the incentive schemes – as the participants go about their own activities and are not ‘bothered’ with additional, potentially intrusive crowdsourcing tasks, they may not require to be explicitly motivated to join.

Validating and aggregating the results produced by the crowd is a core component of every crowdsourcing project. First and foremost, the requester cannot assume that contributions will be usable as they arrive. This is due to many reasons, from subpar skills to bad will to an ill-defined crowdsourcing call. Either way, the requester will have to assess what the crowd created. When microtasks are involved, the hope is that this activity will be carried



out mostly automatically (Quinn & Bederson, 2011). One common approach is redundancy – having multiple members of the crowd undertake the same task and using a weight metric (e.g., majority voting) to identify the answer which is most likely to be correct. More sophisticated solutions have been intensively researched in the crowdsourcing literature. In Ipeirotis, Provost and Wang (2010), for instance, weights take into account the previous performance of workers.

Similar learning techniques are applied to optimize other aspects of the process as well, from assigning tasks to the best skilled or available contributors (Le, Edmonds, Hester, & Biewald, 2010; Raykar & Yu, 2011), to grouping related tasks into bunches (Heymann & Garcia-Molina, 2011) to explore economies of scale and learning effects, to training and providing feedback to improve performance and retention (Dow, Kulkarni, Bunge, Nguyen, Klemmer, & Hartmann, 2011).

## 5.5. How to incentivize

The social fabric of a crowdsourcing endeavour creates a number of challenges for the designers of related platforms and experiments. Let's remember that the requester's aim is to achieve its goal within certain constraints, which can be a certain level of quality, or time, or resources. No matter what flavour of crowdsourcing one would go for, the people contributing will behave according to their own motivation and will react to the incentive mechanisms put in place by the requester. Essentially one can distinguish between three classes of scenarios in this context (Malone et al., 2010): love, glory, and money, mapping more or less to intrinsic motivators, extrinsic motivators, and financial incentives. Love and glory stand for scenarios in which participation is on a voluntary basis and there are no remunerations involved – the crowd finds the tasks enjoyable or rewarding in themselves, or the status that comes with their involvement. Volunteer crowdsourcing has many advantages: it saves costs, and can, at least for a while, attract a large number of contributors. Citizen science projects, Wikipedia, or social good campaigns all stand as manifestations of how powerful and effective it can be. However, it is also difficult to replicate for all types of tasks, especially when these are repetitive, unpleasant, or less important for the potential contributors. Reward models, by contrast, are often easier to control and study, though they are not without challenges (Cuel et al., 2011).

The requester has to decide what to pay for and how much and the behaviour of the crowd will be a function of both (Simperl, Cuel, & Stein, 2013). There is a large body of research in the crowdsourcing community studying optimal pricing strategies (Archak, 2010; Singer & Mittal, 2013) or the interplay between incentives and other factors such as existing social structures (relevant, for instance, when a crowdsourcing project is carried out in an existing social network or in an enterprise) or motivation. Studies such as (Feyisetan, Simperl, van Kleek, & Shadbolt, 2015; Kaufmann, Schulze, & Veit, 2011) have shown that contributors to crowdsourcing platforms are driven by a rich set of reasons and that relying on financial incentives only might miss great opportunities for engagement and performance gains.

An interesting case in studying incentives engineering are games with a purpose and gamification techniques. They are associated with human computation-style of tasks, which tend to be repetitive and not always intellectually stimulating, which are not necessarily rewarding in themselves. Adding game elements to the job is expected to raise motivation, as players take on challenges, receive immediate feedback on their performance, and can compete against others. However, developing a good GWAP is often more tricky than it seems. Some tasks and domains will remain more appealing and accessible to gamers than others – common knowledge subjects, sports, celebrities, some scientific disciplines tend to work better than, say, product catalogues or taxonomies of economic terms. The game narrative is equally important, alongside some knowledge of the task domain that allows the requester to reliably implement core game elements like feedback and levels. Going back to the example we had on publications and data sets, let's say we have an idea for a game to turn the relatively unexciting task of linking the two into something people could imagine doing at high speed and for a longer period of time. Building in something as basic as levels ideally requires a means to distinguish between easier and more challenging data set citation tasks; if this is not the case, one would most likely shorten the time or ask for additional information to make the tasks more difficult. Giving feedback implies that the requester has an idea about the data sets mentioned in publications – a gold standard is needed, which takes time and resources to build. Alternatively one could think about multi-player models with peer assessment (Von Ahn & Dabbish, 2008) or about rewarding participation instead of accuracy in the hope that redundancy in answers will compensate for that. Even so, the players will constantly need additional features to keep them coming back to the game; this is not easy to justify for one-off crowdsourcing projects (Thaler, Simperl, & Wölger, 2012).

## 6. From Crowdsourcing to Social Machines

From the discussion so far it should become clear that there is no single way to crowdsource a task. There are crowdfunding (funding a venture with the help of monetary contributions from the crowd), microtasks (crowdsourcing for routine work broken down into smaller, independent units), macrotasks (closer to classical outsourcing), challenges (competitions targeting grand scientific, technology, business, or social questions), and 'volunteer campaigns' (initiatives seeking ideas and contributions for the public good), and possibly many others, and each potential requester will have to understand what these forms of crowdsourcing mean and the extent to which they would be suitable for their goal. For each of them, there are dozens if not hundreds of service providers to choose from, offering varying level of assistance and support in terms of project set up and operation.

This article has tried to shed light on the richness of this design space, in which there will be many options to consider and a potential requester will have to pick some and possibly use them in combination, while understanding boundary conditions, effectivity, and the impact of incentives. Ideally we are aiming at systems combining both crowd and machine intelligence, which some call social machines; they leverage the best attributes of human capabilities and computing technologies to realize something that meets not only purely utilitarian reasons, but might turn out to be rewards and ethically responsible as well.

A number of prominent examples aside, today we still see a divide between these two forms of intelligence, between conventional IT systems dedicated to data-and computation-intensive tasks, and Web 2.0 sites offering some combination of well-known participatory features, in which user-generated content and the underlying social network evolve dynamically and hand-in-hand. However, as technology becomes more and more ubiquitous, many of the challenges we witness in any domain of our life, business, and society will soon require solutions that rely on both of these axes: a sophisticated combination of data-intensive, complex automation and deep community involvement. This suggests the need for new types of systems to tackle these emerging challenge and for a thorough understanding of the science and engineering of (the continuum of) social machines.

## 7. References

- Archak, N. (2010). Money, glory and cheap talk: analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder. com. In *Proceedings of the 19th international conference on World Wide Web* (pp. 21–30). New York: ACM. doi: 10.1145/1772690.1772694.
- Burke, J.A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., et al. (2006). Participatory sensing. *U.C.L.A. Center for Embedded Network Sensing*, Retrieved March 2, 2015, from: <https://escholarship.org/uc/item/19h777qd>.
- Chesbrough, H. (2003). *Open innovation: the new imperative for creating and profiting from technology*. Boston: Harvard Business School Press.
- Cuel, R., Morozova, O., Rohde, M., Simperl, E., Siorpaes, K., Tokarchuk, O., et al. (2011). Motivation mechanisms for participation in human-driven semantic content creation. *International Journal of Knowledge Engineering and Data Mining*, 1(4), 331–349. Retrieved March 3, 2015, from [http://www.wiwi.uni-siegen.de/wirtschaftsinformatik/paper/2011/cuel-motivation\\_participation\\_2011.pdf](http://www.wiwi.uni-siegen.de/wirtschaftsinformatik/paper/2011/cuel-motivation_participation_2011.pdf).
- Dawson, R., & ByngHall, S. (2012). *Getting results from crowds*. San Francisco, CA: Advanced Human Technologies.
- Dow, S., Kulkarni, A., Bunge, B., Nguyen, T., Klemmer, S., & Hartmann, B. (2011). Shepherding the crowd: managing and providing feedback to crowd workers. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1669–1674). New York: ACM. doi: 10.1145/1979742.1979826.
- Drăgan, L., Luczak-Rösch, M., Simperl, L., Packer, H., Moreau, L., & Berendt, B. (2015). A-posteriori provenance-enabled linking of publications and datasets via crowdsourcing. *D-Lib Magazine*, 21(1/2). Retrieved April 7, 2015, from <http://www.dlib.org/dlib/january15/dragan/01dragan.html>.
- Feyisetan, O., Simperl, E., van Kleek, M., & Shadbolt, N. (2015). Improving paid microtasks through gamification and adaptive furtherance incentives. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 333–343). New York: ACM. doi:10.1145/2736277.2741639.
- Grudin, J. (1994). Computer-supported cooperative work: history and focus. *Computer*, 27(5), 19–26. doi: 10.1109/2.291294.
- Haythornthwaite, C. (2009). Crowds and communities: light and heavyweight models of peer production. In *on* (pp. 1–10). Washington, DC: IEEE. doi: 10.1109/HICSS.2009.137.
- Hendler, J., & Berners-Lee, T. (2010). From the semantic web to social machines: a research challenge for AI on the World Wide Web. *Artificial Intelligence*, 174(2), 156–161. doi: 10.1016/j.artint.2009.11.010.

- Heymann, P., & Garcia-Molina, H. (2011). Turkalytics: analytics for human computation. In *Proceedings of the 20th international conference on World Wide Web* (pp. 477–486). New York: ACM. doi: 10.1145/1963405.1963473.
- Hitchcock, S., Brody, T., Gutteridge, C., Carr, L., Hall, W., Harnad, S., et al. (2002). Open citation linking: the way forward. *D-Lib Magazine*, 8(10). Retrieved April 7, 2015, from <http://www.dlib.org/dlib/october02/hitchcock/10hitchcock.html>.
- Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*, 14(6), 1–4. Retrieved April 7, 2015, from [http://archive.wired.com/wired/archive/14.06/crowds.html?pg=1&topic=crowds&topic\\_set=](http://archive.wired.com/wired/archive/14.06/crowds.html?pg=1&topic=crowds&topic_set=).
- Ipeirotis, P.G., Provost, F., & Wang, J. (2010). Quality management on Amazon Mechanical Turk. In *Proceedings of the ACM SIGKDD workshop on human computation* (pp. 64–67). New York: ACM. Retrieved April 7, 2015, from <http://misrc.csom.umn.edu/workshops/2012/fall/Ipeirotis.pdf>.
- Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker motivation in Crowdsourcing – A study on Mechanical Turk. In *Proceedings of the Seventeenth Americas Conference on Information Systems (AMCIS)* (pp. 1–11). AIS Electronic Library.
- Le, J., Edmonds, A., Hester, V., & Biewald, L. (2010). Ensuring quality in crowdsourced search relevance evaluation: the effects of training question distribution. In *Proceedings of the SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation* (pp. 21–26). New York: ACM. Retrieved April 7, 2015, from <http://ir.ischool.utexas.edu/cse2010/materials/leetal.pdf>.
- Lévy, P. & Bonomo, R. (1999). *Collective intelligence: Mankind's emerging world in cyberspace*. Cambridge, MA: Perseus Publishing.
- Malone, T., Laubacher, R., & Dellarocas, C. (2009). Harnessing crowds: mapping the genome of collective intelligence. MIT Sloan Research Paper No. 4732-09. Retrieved April 7, 2015, from <http://cci.mit.edu/publications/CCIwp2009-01.pdf>.
- Malone, T.W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review* 51(3). 20–31. Retrieved March 2, 2015, from <http://gaius.cbpp.uaa.alaska.edu/afef/CollectiveIntel.pdf>.
- Oleson, D., Sorokin, A., Laughlin, G., Hester, V., Le, J., & Biewald, L. (2011). Programmatic gold: targeted and scalable quality assurance in crowdsourcing. *Human computation: papers from the 2011 AAAI Workshop (WS-11-11)*, pp. 43–48. Retrieved March 3, 2015, from <http://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/download/3995/4267>.
- Parameswaran, M., & Whinston, A.B. (2007). Research issues in social computing. *Journal of the Association for Information Systems*, 8(6), 336–350.

- Quinn, A.J., & Bederson, B.B. (2011). Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1403–1412). New York: ACM. doi: 10.1145/1978942.1979148.
- Raykar, V., & Yu, S. (2011). Ranking annotators for crowdsourced labeling tasks. In *Advances in neural information processing systems 24: 25th Annual Conference on Neural Information* (pp. 1809–1817). Red Hook, NY: Curran Associates. Retrieved April 7, 2015, from <http://papers.nips.cc/paper/4469-ranking-annotators-for-crowdsourced-labeling-tasks.pdf>.
- Simperl, E., Cuel, R., & Stein, M. (2013). Incentive-centric semantic web application engineering. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 3(1), 1–117. doi: 10.2200/S00460ED1V01Y201212WBE004.
- Singer, Y., & Mittal, M. (2013). Pricing mechanisms for crowdsourcing markets. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 1157–1166). New York: ACM. Retrieved April 7, 2015, from [http://www.eecs.harvard.edu/econcs/pubs/Singer\\_www13.pdf](http://www.eecs.harvard.edu/econcs/pubs/Singer_www13.pdf).
- Smart, P., Simperl, E., & Shadbolt, N. (2014). A taxonomic framework for social machines. In D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, & J. Stewart (Eds.), *Social Collective Intelligence*, (pp. 51–85). Cham, Switzerland: Springer.
- Surowiecki, J. (2005). *The wisdom of crowds*. New York: Anchor.
- Tapscott, D. & Williams, A.D. (2008). *Wikinomics: how mass collaboration changes everything*. New York: Penguin.
- Thaler, S., Simperl, E., & Wölger, S. (2012). An experiment in comparing human computation techniques. *Internet Computing, IEEE*, 16(5), 52–58. doi: 10.1109/MIC.2012.67.
- Von Ahn, L. (2009). Human computation. In *Proceedings of the 46th Annual Design Automation Conference (DAC'09)* (pp. 418–419). New York: ACM. doi: 10.1145/1629911.1630023.
- Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, 51(8), 58–67. doi: 10.1145/1378704.1378719.
- Wang, F.-Y., Carley, K.M., Zeng, D., & Mao, W. (2007). Social computing: from social informatics to social intelligence. *Intelligent Systems, IEEE*, 22(2), 79–83. doi: 10.1109/MIS.2007.41.
- Wiggins, A., & Crowston, K. (2011). From conservation to crowdsourcing: a typology of citizen science. In *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS)* (pp. 1–10). New York: IEEE and Computer Society Press. doi: 10.1109/HICSS.2011.207.

## Notes

---

<sup>1</sup> See also <http://googleblog.blogspot.de/2009/09/announcing-project-10100-idea-themes.html>.

<sup>2</sup> See also <https://www.techdirt.com/blog/innovation/articles/20120409/03412518422/why-netflix-never-implemented-algorithm-that-won-netflix-1-millionshtml>.

<sup>3</sup> <http://crowdfunder.com/>

<sup>4</sup> <http://dailycrowdsourc.com/>

<sup>5</sup> <https://www.mturk.com/>

<sup>6</sup> <http://figshare.com/>

<sup>7</sup> [www.researchgate.net/](http://www.researchgate.net/)

<sup>8</sup> [www.mendeley.com/](http://www.mendeley.com/)

<sup>9</sup> <http://www.taverna.org.uk/>

<sup>10</sup> <https://www.historypin.org/>

<sup>11</sup> <https://www.zooniverse.org/>

<sup>12</sup> <http://www.innocentive.com/>

<sup>13</sup> <https://www.kaggle.com/>

<sup>14</sup> <http://polymathprojects.org/>

<sup>15</sup> <http://www.pgdp.net/c/>

<sup>16</sup> <https://www.duolingo.com/>