Crowdsourcing

Classification, costs, benefits, and usage

Rogier Brussee, Michiel Rovers, Harry van Vliet, Dick Swart, Erik Hekman[1]

Rogier.Brussee@hu.nl, Michiel.Rovers@hu.nl, Harry.vanVliet@hu.nl, Dick.Swart@hu.nl, Erik.Hekman@hu.nl

Abstract

The term crowdsourcing was introduced by Jeff Howe (2006). It is the act of a company or organisation to take a function once performed by employees and outsourcing it to an undefined, and usually large, network of people in the form of an open call. As communication tools to organize work have become widely available, and a well-educated global work force has come online, crowdsourcing has become an increasingly important mechanism to organize work. We discuss a categorisation of crowdsourcing, its costs and benefits and several examples. The use of crowdsourcing begins with the question which strategic goal an organisation wants to achieve, and whether the benefits outweigh the costs. We give some recommendations for adopting crowdsourcing. This usually requires a certain amount of restructuring of existing workflows and a willingness to become more open which may or may not be a welcome side effect.

Introduction

Jeff Howe (2006), introduced the term crowdsourcing in Wired Magazine: "[...] crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers." Howe (2008) identifies four developments that have contributed to the rise of crowdsourcing: the increase in the number of well educated, well informed "amateurs" (often using their professional experience); the increasing popularity of open source software as a widely publicised and successful example of collective production (in this case of software); the availability of online communication and publication tools made possible by the web, and the growth of online communities made possible by these online tools.

This paper is based on Van Vliet et al (2013). It is organised as follows. We first categorise different forms of crowdsourcing following earlier work by Oomen & Arroyo (2011). As the organising principle for the classification we use the outcome of the crowdsourcing process. We then focus on the cost and benefits of different forms of crowdsourcing. In particular we discuss the work that is required to organise the crowdsourcing process, and organisational changes that may be required to use it. Whether the benefits outweigh the costs depends on the strategic goals that one wants to attain.

Classification and Examples of Crowdsourcing

Several authors have tried to classify different forms of crowdsourcing. Kozinets, Hemetsberger & Schau (2008) have a typology of consumer communities: *crowds* in which participants actively contribute to a specific problem that they try to solve, *hives* in which some participants contribute more than others but in where the focus is still on particular innovations, *mobs* in which some participants contribute more, but in which the focus is lifestyle and special interests rather than innovation and *swarms* in which participants freely communicate and contribute. The cocreation types of Fronteer (Pater, 2009, Vollens, 2011) distinguish *clubs of experts* working on a challenge posed by an initiator, *coalition of parties* in which experts work on a common problem like the development of standards, *community of kindred spirits* in which a large group cooperates in finding a solution, and the *crowd of people,* in which an initiator openly proposes a specific challenge to the public. Oomen & Arroyo (2011) classify different forms of crowdsourcing based on the result of the crowdsourcing process. Their main context is crowdsourcing in the cultural heritage sector. They propose the following categories: correction/transcription, contextualisation, complementing collections, classifications, co-curation and crowdfunding.

Following the categorisation in van Vliet (2013), the current classification follows in the steps of Oomen & Arroyo (2011). It reorganises some categories and adds additional ones however,

reflecting the broader scope of the current article. The categories are not exclusive: many projects have elements from several categories, although there is usually a main one.

Produce: Produce something

The focus in this category is on producing a tangible result and the process is organized accordingly. In this category we find most of the examples of crowdsourcing. A very well known example is Wikipedia. Open source projects also fall in this category. Less well known examples are "Get a slogan" (www.getaslogan.com) where inventing a slogan for a company is central, and "Threadless" (www.threadless.com) a T-shirt company whose business model is based on leaving the design to the public. "The Johnny Cash project" (www.thejohnnycashproject.com) is an initiative for collecting drawings of Johnny Cash, which are integrated in a music video for his song "Ain't No Grave". Strung together in sequence, it will be a continuously changing, living tribute to the artist. "Artgamelab" (www.sfmoma.org/exhib_events/exhibitions/453) is a project of the San Francisco Museum of Modern Art, which asked conceptual artists, game- and experience designers to come up with ideas for games that can be played in the museum. "Galaxy Zoo" is an astronomy project that calls on volunteers to classify galaxies. It has led to scientific publications on the statistics of galaxies (Lintott et al. 2008) and the discovery of a new class of astronomical objects (Dutch: voorwerp): "Hannies voorwerp" (en.wikipedia.org/wiki/Hanny's_Voorwerp)

Economically important, are the many crowdsourcing labour markets that have sprung up online. The best known is probably the "Mechanical Turk" (www.mturk.com/mturk/welcome). It is a global labour market organised like a "human cloud computing" platform. Requesters post work, like typing-in business cards, together with some (low) reward. Workers from all over the world look for a task they can complete. Local online labour markets for painting and carpentry work (homeadvisor.com, werkspot.nl) or ICT work, graphic design etc. (www.peopleperhour.com, freelancer.nl) are also thriving. Even research and development is crowdsourced: at "innocentive.com" companies propose problems for which they seek intellectual property. Patentable inventions are paid some more sizeable reward.

Much of crowdsourcing on the Internet revolves around content and information. For crowdsourcing initiatives where the tangible outcome is new or enriched information we can classify further using the mental model of the web as a giant database. We then make a distinction between organising, information and manipulating it. Moreover, we distinguish the fundamental operations for databases: create, read/retrieve (i.e. select and find), update and delete (CRUD). Since the analogy goes only so far, we feel there is distinction between adding and correcting information and split the update category into (at least) two. We therefore distinguish:

- Create: create new information or content. Often creation does not start from scratch, but reuses other material. For example, "You Tube" can be considered as a media business

whose business model is based on crowdsourcing content in return for hosting and distributing it, rather than producing or buying it. Like other media businesses such as commercial TV channels, it makes its money from advertising.

- Retrieve: find and select information. Many of the tasks at the Mechanical Turk are finding information on a particular subject.

- Add: add information or content. For example "Flickr the Commons" allows the public to annotate pictures posted by archives and museums with stories, explanations, names etc.

- Correct/Verify: Correct information or content. For example the "Trove" project of the National Library of Australia (www.trove.nla.gov.au) had OCR'ed its collection of historic newspapers and magazines, and asked the public to verify and correct this information.

- Delete: delete information or content. For example You Tube has a non-offensiveness policy, but relies on the public to implement it.

- Systematise: bring order in information or content. For example, "delicious.com" allows its users to tag different websites. It makes these tags available to all its users thereby making the web more accessible.

- Analyse: Analyse information or content. "Galaxy Zoo" can be considered as an example of this type. Some of the tasks at the Mechanical Turk go beyond looking for information to reworking information into a report.

Rate: rate or value something

This category is based on the idea that in many cases the public knows best, because they are involved and have hands on experience. The general public is also less likely to have a commercial interest in giving good (or bad) ratings. A good example is the reviews and ratings at "amazon.com". Other examples are the ratings of hotels and restaurants at "tripadvisor.com". A non standard example is the democratic system, which could be considered as crowdsourcing an evaluation of government.

Facilitate: help make something possible

In this category we find crowdfunding where an open call is made for funding, e.g. on "kickstarter.com" (see also Hekman, Brussee (2013), these proceedings). Projects like SETI@home or Einstein@home have crowdsourced vast amounts of computer time. Finally, many charities, schools, festivals, sport clubs, etc. crowdsource volunteers to help run an event or help in funding.

Publicise: draw attention by crowdsourcing something

This category is disproportionately well known, precisely because the main goal of the crowdsourcing process is to draw attention. A well known example are the many versions of talent scouting on TV (the Voice of/ X-factor/..). While superficially crowdsourcing new talent, they are above all a way to create a television show. Likewise, advertising firms use crowdfunding to create new branded products. However, the brand attention it creates is usually more important than the product itself. For example, a Dutch advertising agency launched a campaign for crowdsourcing a new flavour of Lays chips (http://www.lays.nl/the_battle/). This led to a new flavour that was sold in shops during a year, and the winner is rewarded, but more importantly, to television commercials and advertising for Lays. Likewise, the "My Starbucks Idea" (www.mystarbucksidea.com) seems to be (at least) as much about public relations, as it is about getting good business ideas: of the 150,000 ideas sent in, only 895 were implemented.

Costs and Benefits of Crowdsourcing

The benefits of crowdsourcing workers are diverse. Tokarchuk et al. (2012) distinguish reputation, the expectation of reciprocity, competition, altruism, self-esteem, learning, personal satisfaction, and direct or indirect monetary reward (e.g. in finding jobs). After analysing 250 examples, Malone et al. (2010) find three main motivations: money; love and personal satisfaction, the feeling to be part of a larger team and contributing to a larger good; and fame and recognition. Hertel et al. (2003) finds all mentioned elements in open source projects. It follows that the success of crowdsourcing depends on either a sufficient monetary reward, or the creation of an environment where participants *want* to put in effort.

The most direct benefit for a company or organisation is the work that is done by the crowdsourcing participants. Often this work could not have been done by the organisation itself, because it would have been too expensive, too extensive and time consuming, too boring, or would require knowledge, skills or resources that are not available otherwise. A more immaterial benefit is the possibility to get access to much larger group that is potentially valuable for the organisation, e.g. in this group consensus may be reached that benefits adoption of products or services (this is typical in standardisation processes). The increased exposure of the organisation may lead to a culture of openness of the organisation and may be a PR asset. Finally working with an external group can be a source of personal satisfaction.

The costs of crowdsourcing for an organisation are non-negligible and should be weighed against the cost of doing things in house or using contractor with which a longer standing relation has been built. We distinguish several sources of cost.

- Quality control and time management is difficult when people have no obligation to (continue) work on a project, or to deliver it in way that is useful for the organisation. For

some forms of crowdsourcing, (e.g. the Mechanical Turk model) there is no longstanding relation of trust, although tracking experiences of other work requesters mitigates this. There is also safety in numbers: in a large group there is a much better chance that somebody is available, somebody has made a suitable offer, and that somebody realises an error is made. This does not come for free, however. It requires careful selection, constant quality control, and openness. Constant feedback between the crowdsourcing organisation and crowdsourcers among themselves (even if that exposes weaknesses) can greatly improve quality. The Linux kernel project for example, is famous for its very active mailing list where there is continuous discussion on improvements and relentless code review.

- Crowdsourcing may go together with having to give up a certain amount of authority. If an organization's goal can be reached, at least in part, by an anonymous crowd, it may put the value of the organisation in question.
- For crowdsourcing to work, it usually has to be divided in small manageable pieces that can be worked on by different people. Moreover, this work has to be combined in a larger whole. This is not always possible, and often takes a considerable amount of work, requiring people with sufficient skills to oversee this work.
- Crowdsourcing can require integration of the infrastructure of an organisation with one suitable for crowdsourcing, even to the point that it is best to adopt the latter infrastructure.

Using Crowdsourcing

Holley (2010) has some useful practical recommendations for crowdsourcing in the context of libraries and archives divided in four subcategories. Translated to a more general context they are:

1. **Assignment**: set a clear target that needs several people to achieve. Give an indication of the progress and give feedback on the results;
2. **Systems**: use or set up an infrastructure for the internal organisation of people and to set up, manage and complete the project (e.g. mailing lists, document repositories, social media, existing crowdsourcing platforms, Pay-Pal, etc.) Make sure systems are reliable, easy to use and cheap. Give participants the opportunity to choose their actions.
3. **Work**: create tasks that tickle some interest e.g. because it relates to the interests of a group of people, is related to current events, is playful and/or competitive, has interesting technical challenges, is important for society or their loved ones, allows people to do something back for benefits they received earlier from the crowdsourcing work etc. Alternatively, set a reasonable reward. Make the task easy to fulfil. Create a stimulating and attractive website.

4. **Community**: create an online community of participants. Treat the contribution from every participant visibly and respectfully. When the work is non-trivial, make sure that participants learn from each other and from you.

The Crowdsourcing Canvas (www.battleofconcepts.nl) organises the key success factors of a crowdsourcing project: the strategic goal of crowdsourcing; the work that is presented to the crowd; the required skills of the participants; the motivations of participants and the required infrastructure.

Conclusion

Using crowdsourcing starts with the strategic question which goal one wants to achieve. Crowdsourcing is not suitable for all problems, as it requires being able to divide work in manageable pieces that have to be coordinated and integrated. Using crowdsourcing also means accepting a certain lack of control and therefore requires quality control and cherry picking the good work. It also requires a certain openness and a mechanism to make the work worthwhile, monetary, but often (also) through support, challenge and encouragement. As Holley (2010 p. 15) writes, "If the public is given a high level of trust and responsibility, they will respond with loyalty and commitment". Coping with lack of control and a culture of openness may be the greatest challenges for an organisation.

References

Hertel, G., Niedner S. & Herrmann, S. (2003), Motivation of software developers in Open Source projects: an internet-based survey of contributors to the Linux kernel. *Research Policy, 32 pp. 1159--1177* (avaiable on http://nicomedia.math.upatras.gr/Free-OpenSource/Hertel_etal_MotivationInOSS_Survey.pdf last accessed sept. 29 2013)

Hekman, E., Brussee, R. (2013) Crowdfunding and Online Social Networks, *2nd CARPE conference, Manchester Nov. 4-6, these proceedings*.

Holley, R. (2010) Crowdsourcing: how and why should libraries do it? *D-lib Magazine, 16, ¾, March/April 2010*.

Howe, J. (2006) Crowdsourcing, a definition. *Wired 2, June 2006,* available at http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.htm

Howe, J. (2008) Crowdsourcing. How the power of the crowd is driving the future of business. London: Random House.

Kozinets, R.V., Hemetsberger, A. & Schau, H.J. (2008) The Wisdom of Consumer Crowds, Collective innovation in the age of networked marketing. *Journal of Macromarketing, 28, 4, pp.339--354.*

Lintott, C.J. et al. (2008) Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey *Monthly Notices of the Royal Astronomical Society.* 389, 3. pp. 1179--1189, available at www.arxiv.org/abs/0804.4483

Malone, T.W., Laubacher, R. & Dellarocas, C. (2010), The Collective Intelligence Genome. *MIT Sloan Management Review, 51, 3, pp. 21-31.*

Pater, M. (2009) Co-creation's five guiding principles. Amsterdam:Fronteer, available at www.slideshare.net/fronteer-strategy-5-guiding-principles-110426-presentation)

Oomen, J. & Arroyo, L. (2011) Crowdsourcing in the cultural heritage domain: opportunities and challenges. *C&T '11, proceedings of the 5th international Conference on Communities and Technologies pp. 138--149.* New York: ACM.

Surowiecki, J. (2004), The wisdom of the crowds. Doubleday; Anchor.

Tokarchuk, O. Cuel, R.& Zamarian, M. (2012), Analyzing Crowd Labor and Designing Incentives, for Humans in the Loop. *IEEE internet computing, sept/okt 2012, pp. 45-51.*

Vollens, M. (2011) Winning and Failing Co-Creation Platforms. A benchmark study focussed on Co-Creation & eMobility. Board of Innovation. (available at www.boardofinnovation.com/2011/08/22/winning-failing-co-creation-platforms)

van Vliet, H., Brussee, R. , Swart, D. , Hekman, E. & Rovers, M. Crowdsourcing. Univ of Applied Sciences Utrecht. Available at http://www.crossmedialab.nl/files/crowdsourcing.pdf (in Dutch).