Understanding crowdsourcing projects: A systematic review of tendencies, workflow, and quality management

Fábio R. Assis Neto⁎, Celso A.S. Santos

Federal University of Espírito Santo, Av. Fernando Ferrari, 514 - Goiabeiras, Vitória, ES, 29075-910, Brasil

Abstract

Crowdsourcing is a powerful method that leverages human intelligence to solve problems in the medical, linguistics, machine learning, and in a considerable number of other fields of study. This systematic review aims to understand how crowdsourcing projects are designed and executed in the state-of-the-art, considering the following dimensions: Task execution, quality management, and platform usage. Our results summarized trends of the important aspects of a crowdsourcing project, such as crowd and task types, crowdsourcing platforms, and activities used to manage the quality; we also addressed functions and limitations in traditional crowdsourcing platforms, the definition of a crowdsourcing workflow, and the lack of standardization when developing a crowdsourcing project. As future works, we wish to design a conceptual model that represents the important elements of a crowdsourcing project and their relationships, also, we wish to do further researchers focusing on how the quality mechanisms listed in this work should be used in crowdsourcing projects to get reliable results.

1. Introduction

Crowdsourcing, in short, is a general-purpose problem-solving method, which uses a group of participants willing to help solve a proposed problem (Doan, Ramakrishnan, & Halevy, 2011). This method has proven useful in medicine (Dumitrache, Aroyo, Welty, Sips, & Levass, 2013; Foncubierta Rodríguez & Müller, 2012; Irshad et al., 2017; Maier-Hein et al., 2014; Vernez et al., 2017; Zhai et al., 2013), linguistics (Chowdhury et al., 2014; Enochson & Culbertson, 2015; Lasecki et al., 2013), machine learning (Aigrain et al., 2016; Noronha, Hysen, Zhang, & Gajos, 2011; Roemmele, Archer-McClellan, & Gordon, 2014), software development and testing (Dietl et al., 2012; Dwarakanath et al., 2015; Fava, Shapiro, Osborn, Schäfer, & Whitehead Jr., 2016; Leicht, Knop, Blohm, Müller-Bloch, & Leimeister, 2016; Weidema, López, Naye Baziz, Spanghero, & van der Hoek, 2016), and in a considerable number of other fields of study.

As stated by Yin, Gray, Suri, and Vaughan (2016), the crowdsourcing platforms hide personal attributes and social characteristics of participants. In these platforms, the communication usually is only in one direction from the person asking for a solution (the requester or problem owner) to each participant individually (the worker). Hence, participants commonly are abstracted as merely a black box method to accomplish microtasks associated with the problem solution. However, other authors consider the communication in crowdsourcing as a two-way process, such as stated by Oleson et al. (2011), which asked for feedback from the workers to generate ground truth data (gold units) in their crowdsourcing project. Furthermore, complex tasks require increased communication with the workers. As tasks become more complex, the relationship between the crowdsourcer and workers become more akin to outsourcing than crowdsourcing (Staffelbach et al., 2015).
A crowdsourcing process can be viewed as both collaborative or cooperative. Roschelle and Teasley (1995) defined two statements that try to demonstrate the difference between collaborative and cooperative work:

Collaboration is a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem. (p. 70)

Cooperation is accomplished by the division of labor among participants as an activity where each person is responsible for solving a portion of the problem. (p. 70)

In a crowdsourcing collaborative process, collaborators usually have a broad view of the entire process and frequently rely on the communication and contributions of other participants to realize their work, e.g., Idea Jam\(^1\) and Ushahidi\(^2\) approaches. In another hand, a crowdsourcing cooperative process usually is controlled by a human computation algorithm, in which workers contribute individually, and applied to solve a problem proposed by a crowdsourcer and hidden from the workers, e.g., Mturk.

By following a workflow, each member of the crowd works on a portion or on the entire problem, performing a human computation task which produces useful results towards the solution to the crowdsourcer problem. Planning, coordination, and control of process are defined by a production manager, usually, the crowdsourcer, although other authors argue that the responsibility for workflow design can be shared between the crowd and the crowdsourcer (Kulkarni, Can, & Hartmann, 2012).

Kittur et al. (2013) defined that crowdsourcing workflows are needed once they facilitate decomposing tasks into subtasks, managing the dependencies between subtasks, and assembling the results. However, our view of this concept is broader. Crowdourcing workflow is context-oriented and should be composed of at least one or more tasks sets and one or more quality control activities to guarantee reliable results, which can be executed in parallel, iterative, and sequential ways. Considering our conceptions, it is possible to set quality activities before, during, or after the task execution stages, thus allowing the crowdsourcer to better understand how the quality will be managed during his project. Considering that a workflow can contain multiple task sets, each set is composed of tasks with the same objective, which are executed following a workflow. Thus, our conceptions conform the ideas of Kittur et al. (2013), hence facilitating decomposing a complex task into subtask sets, managing their dependencies, and assembling their results.

Hence, in our conception, a crowdsourcing workflow is a unique entity related to each project. It should represent the planning and coordination done by the crowdsourcer in the entire project, thus, representing a broader view of who will participate, what will be produced, and how it will be done at every stage of the workflow.

Popular studies in crowdsourcing, as in Estellés-Arolas and González-Ladrón-de Guevara (2012), Brabham (2012), and Zhao and Zhu (2014) present the development of crowdsourcing field of study by evaluating known systems and general-purpose platforms, as the Amazon Mechanical Turk (Mturk),\(^3\) CrowdFlower,\(^4\) Innocentive,\(^5\) and Threadless\(^6\).

Classic definitions of the crowdsourcing term usually neglected concepts that we consider crucial for every project, in particular the definition of a crowdsourcing workflow and methods to control the quality of the entire project. Hence, this issues motivated us to review the state-of-the-art of crowdsourcing and try to understand how crowdsourcing projects were executed, not only how systems and platforms works. Thus aiming to validate our assumptions and disseminate new conceptions.

In this review, we extend the work presented in previous crowdsourcing surveys and address essential points to be considered when developing a crowdsourcing project. If popular papers in this domain evaluated crowdsourcing platforms and systems, our view is broader, regards the evaluation of quality management and the essential elements of a crowdsourcing project. Furthermore, we present classifications and trends regarding the crowd, tasks, platforms, task results, products generated, and the quality management in crowdsourcing projects. Thus, one of our goals is to highlight the main concerns of crowdsourcing projects. These concepts provide crowdsourcers and researchers with essential information to successfully design their projects. This paper is, therefore, an up-to-date systematic literature review of crowdsourcing projects that aims to provide detailed information about this new research field. Hence, this review presents the following primary contributions:

- Classifications and tendencies of crowd types, task types, crowdsourcing platforms, and activities used to manage the quality of a crowdsourcing project.
- Observations about the limitations of traditional crowdsourcing platforms.
- Our view of the crowdsourcing workflow, regarding tasks and quality management.

This paper is organized as follows: Section 2 presents state-of-the-art background; Section 3 introduces the methods used in our approach; Section 4 presents the results of our review; Section 5 presents a discussion about all information obtained during the review process. Finally, Section 6 proposes future works and concludes this paper.

---

1 ideajam.io
2 ushahidi.com
3 mturk.com
4 crowdflower.com
5 innocentive.com
6 threadless.com
2. Background

The first report about crowdsourcing related to the information technology field of study was written by Howe (2006), and published by Wired Magazine. This article presents a comparison between an online tool called IStockPhoto, which collects and provides images using crowdsourcing, to the services offered by a professional photographer.

To define crowdsourcing, Howe (2008) proposed the following description:

Crowdsourcing is the act of taking a task traditionally performed by a designated agent (such as an employee or a contractor) and outsourcing it by making an open call to an undefined but large group of people. Crowdsourcing allows the power of the crowd to accomplish tasks that were once the province of just a specialized few. (p. 1)

However, the crowdsourcing field of study grew in popularity, and diverse authors proposed several new definitions for the term. In a renowned publication, Estellés-Arolas and González-Ladrón-de Guevara (2012) reviewed various papers and proposed a unified description:

Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage that what the user has brought to the venture, whose form will depend on the type of activity undertaken. (p. 197)

The work of Hosseini, Phalp, Taylor, and Ali (2014) focused on defining four essential pillars of crowdsourcing. The crowd that consists of workers who take part in a crowdsourcing project; the crowdsourcer which is the entity that plans, coordinates, and controls the crowdsourcing project; the crowdsourcing task consists of the activities to be solved by the workers; and the fourth pillar refers to the crowdsourcing platform, which manages the crowd and the tasks.

Considering the proposed definition, Estellés-Arolas and González-Ladrón-de Guevara (2012) formulated eight characteristics and used them for evaluating crowdsourcing systems. These components are listed below:

1. “There is a clearly defined crowd.”
2. “There exists a task with a clear goal.”
3. “The recompense received by the crowd is clear.”
4. “The crowdsourcer is clearly identified.”
5. “The compensation to be received by the crowdsourcer is clearly defined.”
6. “It is an online assigned process of participative type.”
7. “It uses an open call of variable extent.”
8. “It uses the Internet.” (p. 197)

Alongside the eight characteristics listed above, Zhao and Zhu (2014) presented a conceptual framework, based on the design questions behind collective intelligence, proposed by Malone, Laubacher, and Dellarocas (2010), that leverages important concepts in the crowdsourcing domain. This framework proposes the following key questions to differ this domain from others:

1. “Who is performing the task?”
2. “Why are they doing it?”
3. “How is the task performed?”
4. “What about the ownership and what is being accomplished?”

The first question relates to the composition of a crowd, which can be classified as a group of anonymous people or a group of specific people. The second question is about what incentives the workers to contribute to a crowdsourcing project, and according to Ryan and Deci (2000), these incentives can be intrinsic and reflect the natural human propensity to learn and assimilate, or they can be extrinsic and reflect external control or true self-regulation. The third issue relates how the task is performed, whether it will be divided into microtasks that compound a result by blending all microtasks result (e.g., Mturk), or the task will be realized through competition, selecting and rewarding the best results (e.g., freelancer). Tasks can also be solved by free contribution, allowing the workers to use their creativity and define how they can contribute to the project. The last question is about the product generated, that can belong to the public or private domain (Zhao & Zhu, 2014).

Estellés-Arolas, Navarro-Giner, and González-Ladrón-de Guevara (2015) also conducted a literature review searching for different crowdsourcing typologies. Then, they stated an integrated typology composed of five main types:

---

7 wired.com
8 istockphoto.com
9 freelancer.com
1. Crowdcasting: Competitive crowdsourcing initiatives, where a problem is proposed to the crowd and the worker who solves it first or do it better receive a reward.

2. Crowdcollaboration: Crowdsourcing initiatives in which communication occurs between workers of the crowd.

3. Crowdcontent: The crowdsourcer seeks crowd labor and knowledge to create or find content of various types but not in a competitive way.

4. Crowdfunding: When an individual or an organization seeks for funding from the crowd.

5. Crowdopinion: When the crowdsourcer seeks user opinions and validation about a particular issue or product through votes, comments, tags, or ratings.

As crowdsourcing is a general-purpose method, a standard issue is how to classify tasks in this field of study. To address this issue, Brabham (2012) defined four categories of crowdsourcing tasks according to what problem the crowdsourcer seeks to solve. Knowledge Discovery and Management consist of tasks that find and collect information into a common location and format; Tasks that solve empirical problems are classified as Broadcast Search; Peer-Vetted Creative Production organizes tasks that create and select ideas; Tasks that analyze a significant amount of information are classified as Distributed Human Intelligence Tasks.

Besides the previous categories of Estellés-Arolas et al. (2015), Zhao and Zhu (2014), and Brabham (2012), Geiger, Fielt, Rosemann, and Schader (2012) observed two fundamental dimensions related to crowd participation in crowdsourcing systems. The first dimension consists of systems that seek homogeneous or heterogeneous contributions from the crowd: A homogeneous system values all (valid) contributions equally, i.e., all the inputs are qualitatively equal. In contrast, a heterogeneous system values the contributions according to their qualities, i.e., all inputs are qualitatively different. While the second classifies systems that seek emergent or non-emergent values from these contributions: A non-emergent system generates its product by selecting individual contributions, while an emergent system generates its product from the entirety of inputs and the relationship between them, i.e., combining all contributions and selecting the crowd consensus.

Considering the previous two dimensions, Geiger et al. (2012) categorized the crowdsourcing systems in four types according to the execution of the crowdsourcing process and how its product is generated.

The Crowd Rating type classifies systems that collect massive homogeneous contribution and seeks to generate their product through aggregations and relations between all contributions. The consensus between collaborations provides a collective answer, such as a spectrum of opinions or common assessments and predictions that reflect the wisdom of crowds (Geiger et al., 2012).

Crowd Processing systems collect massive homogeneous contributions and seek non-emergent value that derives directly from the individual contributions. These systems commonly divide a task into micro tasks and merge all individual contributions of one microtask to generate a collective result, e.g., labeling images (Geiger et al., 2012).

Systems that collect heterogeneous contributions and their products derive from aggregations and relationships between all contributions are classified as Crowd Creation systems. Each input is qualitatively different and aggregating these contributions must generate a comprehensive product, e.g., an image database (Geiger et al., 2012).

Crowd Solving consists of systems that collect heterogeneous contributions and seeks non-emergent value that derives directly from the individual contributions. Specific evaluation criteria define the qualitative value of a contribution. Workers on these systems present alternative or complementary solutions to a proposed problem; then the product is generated by evaluating and selecting the best solutions (Geiger et al., 2012).

In the next section, we appropriately evaluate the information exposed in our state-of-the-art background to define important concepts and classifications used for searching, selecting, reviewing, classifying, and gathering data from crowdsourcing papers and projects.

3. Methods

This section presents methods used in this review for filtering and selecting scientific papers that related crowdsourcing projects. These methods are elaborated based on classifications given by the works cited in the previous section.

3.1. Research questions

This study seeks to understand how crowdsourcing projects are reported by the scientific community. To fulfill this purpose we designed the following research questions:

• (RQ1) What are the tendencies related to crowd management, platforms used, and task types in crowdsourcing projects?
• (RQ2) What are the limitations of general-purpose platforms used in crowdsourcing projects?
• (RQ3) How is quality managed in the crowdsourcing projects?

3.2. Selecting crowdsourcing projects

Considering that the characteristics proposed by Estellés-Arolas and González-Ladrón-de Guevara (2012) were defined with the purpose of analyzing crowdsourcing systems, we compared and contrasted these features with the conceptual framework proposed by Zhao and Zhu (2014).

Estellés-Arolas and González-Ladrón-de Guevara (2012) concluded that a crowd is composed of a group of individuals with the
characteristics of size, heterogeneity, and knowledge determined by each project. To classify the crowd used in each project, we defined that when the project provides individual information about workers to the crowdsourcer, it uses a specific crowd, while projects that don’t provide such information about the workers apply an anonymous crowd.

We did not consider that the use of the Internet in crowdsourcing projects is mandatory because it is possible to use other communication means, as we can examine in the work of Gupta, Thies, Cutrell, and Balakrishnan (2012) and in the Ushahidi platform. These platforms use SMS communication, instead of using the Internet in their crowdsourcing process (Ashley, Corbett, Jones, Garside, & Rambaldi, 2009).

According to previous discussions, we choose to use the four following criteria to filter and select crowdsourcing projects in our research:

1. What is the crowd composition? (Anonymous or Specific)
2. What are the incentives offered? (Intrinsic or Extrinsic)
3. How the tasks are performed? (Defined Task, Competitive Task, or Free Contribution)
4. What product is generated at the end of the project?

Different from Brabham (2012) that seeks to classify crowdsourcing tasks according to the problem being solved, we established four categories of tasks, inspired by the classification of crowdsourcing information systems proposed by Geiger et al. (2012). Fig. 1 shows the four categories of crowdsourcing tasks proposed in our approach. This description was used to evidence three essential matters in the tasks: (i) What is provided by the crowdsourcer when designing a task; (ii) how the crowd works in a task; and (iii) what is the outcome of a task.

The four types of crowdsourcing task are described as follows:

1. Object Production: The crowdsourcer gives a task description requesting that the workers produce objects, each worker generates one or more objects, and the outcome is a set of new objects created by the workers.
2. Object for Solution: The crowdsourcer describes a problem and requests workers to generate solutions, the workers individually produce solutions, and the outcome is the best suitable object that solves the given problem.
3. Object Processing: The crowdsourcer gives a set of N objects to be processed by the crowd, the workers process the objects according to the task description, and the outcome can be a set of objects edited by the workers or the given objects alongside a set of new metadata generated by the crowd.
4. Object Evaluation: The crowdsourcer gives a set of N objects to be evaluated by the crowd, the workers assess the objects according to the task description, and the outcome will be the given objects alongside a set of evaluations related to these objects.

---

Fig. 1. The four types of crowdsourcing tasks.
We divided these four task types into two dimensions:

1. Generation: This dimension classifies task types that the crowdsourcer gives a task description and request that the workers generate objects or solutions for a proposed problem. Each object created in these tasks represents the work of a single worker.
2. Improvement: This dimension classifies task types that the crowdsourcer gives a set of N objects alongside the task description. Each object processed or evaluated in these tasks represents the work of X workers.

3.3. Query string

Aiming to find papers that approach platforms, systems, or applications in crowdsourcing and then highlight articles that focus on planning and applying crowdsourcing projects, we formulate a query string that includes variations of the term crowdsourcing along with systems and platforms that include the terms jobs or tasks. We excluded from our string the terms crowdfunding, crowdsensing, and papers that focus reviews and surveys. Therefore, our query string is defined below:

\[(\text{crowdsourc}^*) \text{ AND (platform OR software OR application) AND (job OR task)} \]
\[\text{NOT (crowdsensing OR crowdfunding OR survey OR review)}\]

This query string was executed on the following digital libraries: (i) ACM Digital Library. (ii) Engineering Village. (iii) IEEE Xplore. (iv) Science Direct. (v) Scopus. These libraries were selected because of their relevance to the field of computer science, and their content is accessible to the authors of this study.

Papers are selected as the result of a match between the query string and their title/abstract/keywords, this research was realized until Oct. 2016. The query string was adapted according to the requirements of each digital library.

3.4. Inclusion and exclusion criteria

Papers that fit the following criteria were selected to be reviewed:

- (INC1) Papers that mainly focus crowdsourcing.
  Rationale: A considerable number of articles use crowdsourcing to compare results or to complement purposes on other fields of study.
- (INC2) Papers that present planning and execution of crowdsourcing projects.
  Rationale: Evaluating crowdsourcing projects allows a broad view of the crowdsourcing process.
- (INC3) Papers that were written in English or Portuguese.
  Rationale: For reasons of feasibility, articles written in another language were excluded.

Papers that fit the following criteria were excluded:

- (EXC1) Short Papers.
  Rationale: We considered that the papers should have at least four pages to be suitable for our review process.
- (EXC2) The four criteria presented in the Section 3.2.
  Rationale: Considering that these criteria are essential to identify a crowdsourcing project, we excluded papers that did not present sufficient information to report each of them.
- (EXC3) Crowdsourcing projects inexistent or poorly defined.
  Rationale: The primary focus of our study is understand the crowdsourcing process as a whole. Therefore, it’s essential to evaluate well-defined projects.
- (EXC4) Similar papers.
  Rationale: When two or more papers have similar authors, crowdsourcing projects, and results, we included the most detailed article.

3.5. Selecting papers

Papers were selected and organized using the Mendeley\(^{10}\) tool, that was chosen because it’s useful for organizing, annotating and filtering papers. After executing our query string on the selected digital libraries and removing duplicated papers, call for papers, and proceeding descriptions we started our study with the following numbers:

Table 1 introduces the 947 papers that composed our first stage of reviews, which consisted of comparing the defined inclusion and exclusion criteria with the information obtained by reading the title, abstract, and keywords of selected papers. This stage resulted in 813 excluded papers and 133 papers selected for the second stage of reviews, which applies inclusion and exclusion criteria on the information obtained reading the articles entirely.

The second stage of reviews resulted in 62 selected papers that presented 66 crowdsourcing projects for further analysis. It is noteworthy that some papers presented multiple crowdsourcing projects, and we considered all the projects that offer different

\(^{10}\) mendeley.com
classifications.

The entire searching and selecting papers process was repeated in May 2017 to update our study and tendencies. Disregarding all papers found in our first search process, removing duplicates, and call for papers, our second search process found 153 new articles. After applying inclusion and exclusion criteria in the title, abstract, and keywords, 31 of 153 new papers were selected for the second stage of reviews. In the second stage of reviews, we read and evaluated the 31 articles, and after this step, we selected ten papers that presented well-defined crowdsourcing projects to include in our study. We also published all references found in our query process in a Mendeley Dataset.11

Unifying the two search process mentioned above, we gathered and evaluated data of 76 crowdsourcing projects found in 72 papers.

3.6. Data gathering

To gather relevant data about the selected papers, we classified the crowdsourcing projects using the four criteria mentioned in Section 3.2 plus the three following categories:

1. What is the purpose of the platform used in the project? (General-purpose or Specific-purpose)
2. What type of crowdsourcing task does the project apply? (Object processing, Object creation, Object evaluation, Object for solution)
3. What activities are responsible for managing quality during the crowdsourcing process?

After gathering data by answering the questions addressed above in the 76 crowdsourcing projects, we observed tendencies and classifications that are presented in the next section.

4. Results

4.1. What is the crowd composition?

Our study indicated 50 crowdsourcing projects that the crowdsourcer doesn’t store information about the workers, classified as an anonymous crowd, and 26 projects in which the crowdsourcer stores individual information about the workers, classified as a specific crowd.

Regarding the crowdsourcing projects that approached an anonymous crowd, 74% used general-purpose platforms (e.g. Mturk and CrowdFlower). However, 73,07% of the projects that approached a specific crowd used specific-purpose platforms.

It’s noticeable that general-purpose platforms are likely used for managing an anonymous crowd. General-purpose platforms typically manage several crowdsourcing projects, distributing tasks to anonymous workers and collecting results. If a crowdsourcer wants to store information about the workers, it’s possible to control a crowd by developing a specific platform that supplies his needs, defining then a specific-purpose platform.

Table 2 indicates that most of the crowdsourcing projects that manage an anonymous crowd seek to perform object processing (e.g. image segmentation, video annotation, text recognition). Therefore, the lack of information about the workers, that defines the pseudo-anonymity of the crowd (Chandler & Kapelner, 2013; Chowdhury et al., 2014; Kim, Li, chul Kwon, & Yi, 2011), limits the execution of complex tasks by anonymous workers. Lack of information about the crowd hinders crowd management, quality management, and evidence demographic limitations (Kim et al., 2011).

Regarding the projects that approach specific crowds, even if 69,23% of the cases seek to perform object processing, studies similar to Feng et al. (2016) are viable. Their study presents the development of an educational game that aims to provide knowledge to students while they analyze biomedical images. Specific crowds also allow projects to be performed in laboratories, university campus, and specific geographic locations, as presented in the following papers: (Figueroa Salas, Adzic, Kalva, & Shah, 2013; Gupta et al., 2012; Kermanidis, Maragoudakis, & Vosinakis, 2015; Lasecki et al., 2012; Tuite, Snively, Hsiao, Tabing, & Popović, 2011).

Storing information about workers enables the management of a crowd composed of specialists in specific knowledge fields, as

---

11 https://doi.org/10.17632/z5mrss39v7.1
4.2. What are the incentives offered?

Hosseini, Shahri, Phalp, Taylor, and Ali (2015) stated that incentives can be intrinsic or extrinsic. Intrinsic motivations are a feature within the crowd, while the extrinsic ones come from the crowdsourcer, in the form of financial incentives or other forms. They also identified five types of intrinsic incentives: (i) Knowledge sharing; (ii) Love of community; (iii) Personal skills development; (iv) Self-esteem; and (v) Volunteering. During our classification, we summarized the incentives found within categories that we found relevant to calculate our tendencies.

We observed that 51 of 76 reviewed projects presented extrinsic incentives, while 21 projects used intrinsic incentives and four projects approached multiples motivations, that incentivized workers intrinsically and extrinsically.

Considering the information shown in Table 3, it’s noticeable the predominance of monetary incentives, such as payments for each task done, raffles, and rewards to the best contributors. Financial incentives, mainly used in commercial platforms such as the Mturk, have the advantage of attracting large amounts of people to compose a crowd (Cao, Chen, & Jagadish, 2014; Kim, Kim, Yang, Ryu, & Wohn, 2014; Kim et al., 2011). Nevertheless, monetary incentives can be harmful to a crowdsourcing project, attracting malicious workers that only seeks to obtain financial rewards without caring about the work done (Cao et al., 2014; Kim et al., 2014). Bénabou and Tirole (2003) stated that monetary rewards also can undermine intrinsic incentives and become negative reinforcers once they are removed.

Access incentives are presented in papers which use the CAPTCHA approach (von Ahn, 2005). This method aims to prevent automatic access to websites by giving tests that only humans can solve. Using access as crowdsourcing incentive means that the workers will need to solve tasks in exchange to access specific sites. By solving these tasks the workers are contributing to a crowdsourcing project, as presented in the studies of Law and Ahn (2011), Hillen and Höflé (2015), and Kim et al. (2014).

Another extrinsic incentive is presented in the work of Feng et al. (2016), which motivates workers by providing school supplies to the students that participated in their crowdsourcing experiment. Table 4 shows that 60% of the projects that use intrinsic incentives approached entertainment as motivation for the crowd. This incentive is defined by the term Games-With-a-Purpose (GWAP) attributed by von Ahn and Dabbish (2008), which addresses games developed with the purpose of collecting human intelligence through tasks that entertain the player (von Ahn & Dabbish, 2008; Tuite et al., 2011).

Social incentives are exemplified in the study of Vukovic, Salapura, and Rajagopal (2013), which publicly acknowledge the work done in the crowdsourcing tasks. Considering that the workers in this project belong to the same type of business, acquiring acknowledgment in this group is a viable incentive to execute tasks.

Projects that use altruism as incentive seek to mobilize people to compose a crowd and work to benefit someone else. Studies found include natural disasters, as in the work of Hester, Shaw, and Biwald (2010), which describes the effort to categorize and locate information received by SMS sent from people in an emergency during the earthquake that happened in Haiti, 2010. In this same natural disaster, the study of Zhai, Kijewski-Correia, Hachen, and Maday (2012) seeks to identify damage done in buildings and structures. The work of Huyhn et al. (2014) has the same objective as the study of Zhai et al., but was executed during the earthquakes and tsunamis that happened in Japan, 2011.

Learning incentives are exemplified by the study of Dontcheva et al. (2014), which presents a crowdsourcing platform built around the Adobe Photoshop platform. This crowdsourcing system offers advanced tutorials of image editing and provides training tasks to the workers, and then it’s possible to generate a database of edited images by combining the results of these tasks.

The studies of Feng et al. (2016) and Tuite et al. (2011) exemplifies the multiple incentives approach, by presenting entertainment incentives as the primary incentive and offering extrinsic incentives, school supplies, and financial incentives respectively, to the best contributors.

4.3. How the tasks are performed?

Observing the crowdsourcing projects reviewed, 71 used the defined task model, four cases approached the free contribution approach in the studies of Chowdhury et al. (2014; Dumitrache et al., 2013; Dwarakanath et al., 2015; Foncubierta Rodríguez & Müller, 2012; Huyhn, Eguchi, Lin, & Eguchi, 2014; McAllister Byun, Halpin, & Szeredi, 2015). It’s also possible to assemble users from platforms that aren’t specifically developed for crowdsourcing means, such as the Facebook and Adobe Photoshop, presented in the following studies: (Banks, Rafter, & Smyth, 2015; Dontcheva, Morris, Brandt, & Gerber, 2014; Sabou, Scharl, & Föls, 2013).

Table 2

<table>
<thead>
<tr>
<th>Tendencies in projects featuring anonymous and specific crowds.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crowd members</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Anonymous</td>
</tr>
<tr>
<td>Specific</td>
</tr>
</tbody>
</table>

12 facebook.com
13 adobe.com/photoshop
model, and one focused on the competition method.

Table 5 presents the processing tasks as leading trends, which consists of tasks that request the workers to evaluate an object and produce something, like annotations and markings. Object processing includes combining an object provided by the crowdsourcer with task results produced by workers, thus generating new objects or metadata to complement given objects. This approach was detailed in Section 3.2.

Image processing is exemplified in studies that produced annotations (Foncubierta Rodríguez & Müller, 2012; Redi et al., 2013) or markings (Della Mea, Maddalena, Mizzaro, Machin, & Beltrami, 2014; Estes et al., 2016). Marking images differs from annotating images because the markings are grouped to the visual content of an image, while annotations are independent because they use information contained in the image to generate additional content, such as metadata and text transcription.

Regarding video processing, the work of Sulser, Giangreco, and Schuldt (2014), and Nguyen-Dinh, Waldburger, Roggen, and Tröster (2013) approached video annotation according to events that occurred during the video. The crowdsourcing project presented in Masiar and Simko (2015) seeks to collect video metadata by using the crowd to search and compare the search results. In Di Salvo, Spampinato, and Giordano (2016) workers segmented videos by playing a game about marine species. Figuerola Salas et al. (2013) uses human intelligence to compare video quality.

As for text processing, it’s noticeable in the studies of Zhai et al. (2013) and Dumitrache et al. (2013) the usefulness of crowdsourcing to recognize medical text and combine them to generate reliable databases. The work of Mellebeek et al. (2010) appeals to the crowd for evaluating opinions about car brands, seeking to train an opinion mining system.

Evaluating tasks that produce audio processing, the work of Lasecki et al. (2012) presents an approach for transcribing and subtitling audios on demand. The study of McAllister Byun et al. (2015) proposes to evaluate speech using crowdsourcing. McNaney et al. (2016) describes a mobile app for people with Parkinson’s condition to manage their dialogue, receiving feedback generated by a crowdsourcing project.

Crowdsourcing is a versatile approach when considering tasks and their results, therefore, a remarkable number of reviewed projects differ from the classifications presented above, representing then the others category with 23% of the sampling exposed on Table 5.

Among the others category, three projects used crowdsourcing to verify software. Weidema et al. (2016) formulated tasks that aim to collect alternative design solutions to solve user interface problems, while in Pastore, Mariani, and Fraser (2013) is proposed to evaluate the source code and behavior of a program, aiming to find bugs in the software.

The crowdsourcing project presented in To, Geraldes, Shahabi, Kim, and Prendinger (2016) provides an approach that the workers should visit selected landmarks, in a defined geographic region, and register how the landmark influences their mood. This project generates as a product one map with landmarks and the feeling related with each.

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tendencies of the defined tasks category.</td>
</tr>
<tr>
<td>Defined tasks</td>
</tr>
<tr>
<td>Audio processing</td>
</tr>
<tr>
<td>Image processing</td>
</tr>
<tr>
<td>Text processing</td>
</tr>
<tr>
<td>Video processing</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>total</td>
</tr>
</tbody>
</table>

Table 3
Incentive Incidence
---
Access | 02 (03,64%) |
Monetary | 51 (92,73%) |
Others | 02 (03,64%) |
Total | 55 (100,00%) |

Table 4
Incentive Incidence
---
Altruism | 05 (20%) |
Entertainment | 15 (60%) |
Learning | 02 (08%) |
Social | 03 (12%) |
Total | 25 (100%) |
Vukovic et al. (2013) developed a project that uses the intelligence of a specific crowd to compose technical documents. The work of Banks et al. (2015) proposed a game developed on the Facebook platform, that allows the users to recommend movies to their friends, collecting recommendation data afterwards.

Free contribution tasks allow the crowd to use their creativity during the crowdsourcing process. The studies of Tuite et al. (2011) and Bayas et al. (2016) allow the workers to take and send pictures related to a defined locale.

Lasecki et al. (2013) requests that workers answer questions in a chat system with their own words. All workers of this project will vote, and the answer with most votes will be shown to the user that asked the question.

In competitive tasks, as the name suggests, the workers compete by solving and submitting tasks to the crowdsourcer, who selects and reward the solution that best fits his requests, as approached in the software development website TopCoder.\textsuperscript{14} Dwarakanath et al. (2015) presented a methodology that manages a software development lifecycle, with the code implementation part produced with competitive tasks. Competition is useful to find specialists capable of producing reliable and complex results. However, skilled workers that work in competitive crowdsourcing platforms frequently seeks better monetary rewards than those offered by microtask platforms (Dontcheva et al., 2014).

The second dimension presented in the work of Geiger, Seedorf, Schulze, Nickerson, and Schader (2011) can be an interesting way to classify how crowdsourcing tasks are performed. This dimension, called accessibility of peer contributions, indicates if workers can access each other’s contributions. The four characteristics of this dimension are: (i) None, which means that workers cannot see each other’s contributions; (ii) View, in which all contributions are visible to any potential worker participating in the project; (iii) Assess, in which workers can vote, rate, and comment on other contributions; lastly, the (iv) Modify characteristic, which allows workers to modify and even delete each other’s contributions.

4.4. What product is generated at the end of the project?

Table 6 shows the classification of products generated by crowdsourcing projects, pointing out that crowdsourcing product can differ to the result of a paper. This statement is important because crowdsourcing is considerably used to produce media databases and these databases are used to support systems and train artificial intelligence, as presented in Altmeyer, Lessel, and Krüger (2016), Maier-Hein et al. (2014).

Verification of systems is exemplified in the works of Dietl et al. (2012) and Fava et al. (2016), which proposes to verify systems by using crowd intelligence in GWAP.

The study of Lasecki et al. (2013) presents a chat system powered by crowdsourcing. The workers in this project are monetarily rewarded by voting on the best answer or by producing the most voted answer.

The work of Tuite et al. (2011) developed a competitive game that motivated University of Washington students to take pictures of defined buildings in the university. As a result of this project, 3D building models were generated from the pictures taken by the workers, available at the Photo City website.\textsuperscript{15}

4.5. What is the purpose of the platform used in the project?

This study focuses on understanding how the crowdsourcing process is executed nowadays, therefore, is important to evaluate general-purpose platforms and their limitations. Reviewing crowdsourcing projects allowed us to perceive the positive and negative sides of these platforms, due to papers that developed specific-purpose platforms to execute crowdsourcing projects, which frequently described the limitations in general-purpose platforms that led them to develop a specific platform.

Among the 76 reviewed projects, 44 of them were managed by general-purpose crowdsourcing platforms. Table 7 illustrates that the Mturk platform shows greater popularity between these platforms. This platform is capable of managing the development, execution, and validation of crowdsourcing tasks quickly and at low cost, due to numerous subscribed workers available to solve tasks (Alonso & Mizzaro, 2012; McAllister Byun et al., 2015; Yang et al., 2010).

The Mturk allows the crowdsourcer to manage crowd workers, enabling them to limit worker participation based on previous work done and acceptance, also based on geographic location of the workers, and by designing qualification tests that the workers should pass to be allowed to work in a project (Alonso & Mizzaro, 2012; Can, Odobez, & Gatica-Perez, 2014; Chandler & Kapelner, 2013; McAllister Byun et al., 2015; Yang et al., 2010). However, our study points out the following limitations when managing a crowd on this platform: (i) Lack of information and control about the workers (Chandler & Kapelner, 2013; Chowdhury et al., 2014; Kim et al., 2011); (ii) crowd participation is geographically limited (Chowdhury et al., 2014; Della Mea et al., 2014; Ross, Zaldivar, Irani, & Tomlinson, 2010); (iii) worker’s incentive is inflexible (Cao et al., 2014; Kim et al., 2014).

Another limitation found on the Mturk is the difficulty in assembling some tasks, forcing crowdsourcers to host their application on a separate server and post a link in the tasks (Kim et al., 2011; McAllister Byun et al., 2015). Mturk is also limited in the qualification tests and validation mechanisms provided (Alonso & Mizzaro, 2012; Ashikawa, Kawamura, & Ohsuga, 2015; Kim et al., 2011).

The CrowdFlower platform differs from Mturk in two main points, the first is a possibility to use a proper markup language along with HTML, CSS, and Javascript to develop tasks in this platform (Della Mea et al., 2014; Foncubierta Rodríguez & Müller, 2012). The

\textsuperscript{14} topcoder.com

\textsuperscript{15} photocitygame.com
second point is being able to aggregate diverse crowdsourcing platforms as task channels, including Mturk, thus allowing the
crowdsourcer to select in what channels his task will be published (Figuerola Salas et al., 2013; Foncubierta Rodríguez & Müller,
2012).

Unexpectedly, the Facebook social network was used as a general-purpose platform. This platform provides an API for devel-
opment of games and applications, besides providing information about potential workers and their friends. The statements above
allows crowdsourcers to manage a crowdsourcing project on this platform, using entertainment as intrinsic incentive (Banks et al.,
2015; Sabou et al., 2013). It’s noteworthy that the Facebook platform can be used to collect crowdsourcing votes through its comment
system.

The Freelancer platform presents the competitive crowdsourcing approach (Dwarakanath et al., 2015). However, the platform
developed by Brambilla, Ceri, Mauri, and Volonterio (2014) offers an approach that allows the crowdsourcer to modify the execution
of a crowdsourcing project by changing between anonymous and specific crowd during the process.

Microworkers (Redi et al., 2013; Sulser et al., 2014), C-SATS (Vernez et al., 2017), Crowdtesting (Leicht et al., 2016), and oDesk
(Leano, Wang, & Sarma, 2016) are other general-purpose platforms used in the reviewed projects.

Platforms developed to manage a specific experiment were considered as specific-purpose platforms, presented in 32 of the 76
cases. Specific-purpose platforms were eventually developed to supply specific limitations of general-purpose platforms, as presented in
the studies of Brambilla et al. (2014) and Castano, Ferrara, Genta, and Montanelli (2016), which developed platforms that allow a
flexible crowd management. Another important point is that 62,5% of projects that used specific platforms approached intrinsic
incentives, while only 11,36% of projects in general-purpose platforms used them as motivation.

Considering projects in which the crowdsourcer stored information about the workers, 59,37% of projects in specific platform
managed a specific crowd, while only 15,90% of projects in general-purpose platforms approached the specific crowd model.

Table 8 shows that crowdsourcing is mainly used to process objects, disregarding the platform used. The table above also informs
that general-purpose platforms are prone to execute object evaluation and problem-solving approaches, while object production tasks
were only performed in specific-purpose platforms.

Some of the reviewed projects presented the development of a crowdsourcing platform that seeks manage tasks and supply
limitations, found in general-purpose platforms, but still uses these systems to select and control the crowd due to the abundant
number of workers subscribed on these platforms. General-purpose platforms allow the crowdsourcer to design the called external
HITs, which consists of external tasks hosted on another platform, that are provided to the workers through links and access codes, as
presented in the studies of McAllister Byun et al. (2015), and Nguyen-Dinh et al. (2013).

<table>
<thead>
<tr>
<th>Table 8</th>
<th>The incidence of task types in crowdsourcing platform types.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform type</td>
<td>Object processing</td>
</tr>
<tr>
<td>General</td>
<td>36 (82%)</td>
</tr>
<tr>
<td>Specific</td>
<td>23 (72%)</td>
</tr>
</tbody>
</table>
4.6. Quality management in crowdsourcing

Crowdsourcing projects frequently have problems with guaranteeing reliable results to the crowdsourcer (Kim et al., 2011), including cases that provide monetary rewards (Altmeyer et al., 2016; Cao et al., 2014). Therefore, to guarantee quality in the crowdsourcing project and avoid wasting resources, the use of mechanisms to control the quality of a crowdsourcing project is necessary.

Only 10 of 76 crowdsourcing projects reviewed didn’t specify how the quality is managed in their crowdsourcing process. By analyzing the quality management of the 66 remaining studies, we observed the following tendencies: (i) Activities that occur before the crowdsourcing process (Pre-task); (ii) techniques applied during the task execution (During-task); (iii) mechanisms used after the task execution (Post-task).

Table 9 presents the activities used in the pre-task stage, present in 34 of 76 projects, and they are responsible for training and filtering workers, simplifying task execution, and preprocessing objects used to create tasks. However, these methods don’t grant the quality of a product generated by crowdsourcing, because these activities occur before task execution and it’s impossible to predict how the workers will solve the task. Therefore, these techniques help to avoid malicious workers and poorly done work (Alonso & Mizzaro, 2012; Chowdhury et al., 2014; Nguyen-Dinh et al., 2013; Tavares, Mourão, & Magalhaes, 2013).

A way to qualify workers is to insert a mandatory training stage before task execution. This stage normally consists of a tutorial or video that aims to instruct and train workers for executing the crowdsourcing tasks in a project, as presented in the following studies (Dontcheva et al., 2014; Feng et al., 2016; Huynh et al., 2014; Lasecki et al., 2013).

Qualification tests mean to insert a mandatory test that a worker must pass to work on a project. The result of this test defines the ability of the workers to solve the proposed tasks, as approached in Estes et al. (2016), Irshad et al. (2015), Weidema et al. (2016).

Preprocessing consists of the use of algorithms and procedures in the objects provided by the crowdsourcer, aiming to design tasks that can be easily solved by the workers. The work of Dumitrache et al. (2013) preprocesses the objects to add metadata and generate automatic annotations, thus simplifying the tasks provided to the crowd. Soleymani (2015) preprocessed an image database, removing content that is irrelevant to his project.

Worker accuracy is a technique used mainly in general-purpose platforms, because these platforms store execution data of subscribed workers, considering the acceptance of previous work done and accounting for a percentage of approval. This mechanism limits the participation in a project by requiring that the workers should have a defined or higher rate of acceptance (Ashikawa et al., 2015; Can et al., 2014; Yang et al., 2010).

During-task quality management was used in 21 of 76 projects, and 52% of these projects presented a method called task design, which consists of rules inserted at the moment of task execution, aiming to exclude task results that diverge from defined standards. Della Mea et al. (2014) defined that the workers should recognize a minimum number of cells in each image to be considered as a valid task. The work of Zhai et al. (2012) verified the time that a worker spent on a task and compares this time with the average time spent by other workers, thus defining if the worker is executing the task with proper care.

The mechanism called gold standard was used in 62% of the projects that approached during-task quality management, and this technique merges ground truth questions with crowdsourcing tasks. Workers that gives the wrong answers to the ground truth questions were considered unable to respond to a crowdsourcing task correctly, and their results are discarded. This technique was used in the following studies (Hillen & Höfle, 2015; Tavares et al., 2013; Zhai et al., 2013).

Table 10 presents the post-task quality management, used in 57 of 76 projects. These methods are capable of unifying multiple task results and generates a reliable result and also filter undesirable task results after the task execution.

The majority decision, or majority voting, is the most used activity for quality management, present in 41 crowdsourcing projects.

---

**Table 9**

<table>
<thead>
<tr>
<th>Pre-task mechanisms</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Processing</td>
<td>11</td>
</tr>
<tr>
<td>Qualification Tests</td>
<td>12</td>
</tr>
<tr>
<td>Training</td>
<td>15</td>
</tr>
<tr>
<td>Worker Accuracy</td>
<td>08</td>
</tr>
<tr>
<td>Total Incidence</td>
<td>34</td>
</tr>
</tbody>
</table>

**Table 10**

<table>
<thead>
<tr>
<th>Post-task mechanisms</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority Decision</td>
<td>41</td>
</tr>
<tr>
<td>Post-Processing</td>
<td>11</td>
</tr>
<tr>
<td>Review</td>
<td>12</td>
</tr>
<tr>
<td>Subtasks</td>
<td>06</td>
</tr>
<tr>
<td>Total Incidence</td>
<td>57</td>
</tr>
</tbody>
</table>
This technique is frequently used in the microtask approach and consist in the replication of one task in a defined number of identical tasks. After the replicated tasks are solved, the results are aggregated, and the crowd consensus is considered as the task results, as presented in the studies of Irshad et al. (2015), Lasecki et al. (2012), Loni et al. (2014), Tuite et al. (2011), Sabou et al. (2013).

The review method consists of manually reviewing the work done by the crowd, thus deciding the acceptance of the work. This activity is approached in the following papers: (Dontcheva et al., 2014; Dumitrache et al., 2013; Lasecki et al., 2012).

The Post-processing technique consists of optimizing the results obtained in a crowdsourcing project, usually with computational aid. The studies of Hillen and Höflé (2015) and Taborsky, Allen, Blanton, Jain, and Klare (2015) used post-processing to combine the image markups provided by the crowd. Tuite et al. (2011) unified images sent by the crowd to generate 3D models of buildings.

The subtasks mechanism is used to evaluate work done by a crowd. This technique consists of generating new crowdsourcing tasks and request workers to verify objects produced by a crowdsourcing project, as presented in Roemmele et al. (2014), Ramchurn, Huynh, Venanzi, and Shi (2013).

5. Discussion

This section presents the contribution of our study to the state-of-the-art in crowdsourcing, and important points when planning a project in this field of study. Therefore, we related the information obtained by reviewing the 72 papers approached in this study with the four pillars of crowdsourcing presented in Hosseini et al. (2014).

Considering that the crowdsourcer, defined as the first crowdsourcing pillar, manages each project in a unique way, we did not have sufficient information to discuss this entity. Therefore, in the following sections, we discuss issues regarding the crowd, the tasks, and the platform of the reported crowdsourcing projects.

5.1. About the crowd

The crowdsourcer should understand what the crowd will produce in his project. This statement is essential to know if the workers will need specific skills to solve the tasks and if these skills can be taught to unspecialized workers. Dontcheva et al. (2014) applied a mandatory training stage that taught image editing skills to the workers before applying their crowdsourcing tasks. If the necessary skills are too complex to be taught, the crowdsourcer needs to hire workers that have previous knowledge requested to solve the tasks. Dwarakanath et al. (2015) hired workers that already have software development skills.

A common concept of economics is that individuals respond to incentives (Bénabou & Tirole, 2003). Thus, it is also important that the crowdsourcer defines what incentives will be provided to the workers. The incentives offered influence directly in the worker’s interest to participate in a project because most workers who works with others is concerned with motivation, facing the question of how much motivation those others, or oneself, has for a task (Ryan & Deci, 2000).

Section 4.2 presents the incentives used in the crowdsourcing projects reviewed in our study. The costs of monetary rewards and volume of contributions are not approached in our review because this information was not present in all of the reviewed projects.

5.2. About the crowdsourcing task

As aforementioned, crowdsourcing tasks must be linked with incentive mechanisms, which should reward workers according to the incentives established by the crowdsourcer. The crowdsourcer should pay attention if the designed task requires assets, which are resources needed to solve one task. The project reported in Dontcheva et al. (2014) defined that only workers with access to the Adobe Photoshop platform could participate in it. This example shows that is important to understand if a task will require assets because these can limit worker participation in a project.

Gupta et al. (2012) evidenced the importance to associate assets needed with a task. This study selected workers that didn’t have access to computers, smartphones, and the Internet, that are assets required to participate in most of the crowdsourcing projects. This project managed and distributed crowdsourcing tasks via SMS and observed that some workers don’t have access to the standard assets needed to solve crowdsourcing tasks. The crowdsourcer of this project found potential workers that cannot subscribe to general-purpose platforms.

An important concept is that a project should define activities to manage quality in the crowdsourcing process, which were presented in Section 4.6 of this review. The management of quality in the task execution process is critical for saving resources and guaranteeing quality in the final product.

By reviewing crowdsourcing projects, we observed that the crowdsourcer must define how the tasks and quality management activities will interact, thus designing a workflow for the project.

We also perceived that the dominant crowdsourcing infrastructure is the workflow, which decomposes complex goals into small tasks (Retelny, Bernstein, & Valentine, 2017). Zhang et al. (2012) defined that a workflow should manage the inputs and outputs of multiple simple task modules, and the union of these modules will result in the product of a complex task. Furthermore, Hetmank (2013) determined that a crowdsourcing workflow must manage latency, work price, worker quality, and contribution quality.

Considering the definitions above and reviewing the crowdsourcing projects, we defined that a project workflow is unique and context-oriented. Therefore, it is possible to define that all crowdsourcing workflows must have one or more task module and one or more quality management module. Otherwise, it’s impossible to generate reliable results. To exemplify the execution of the workflow.
of a crowdsourcing project, Fig. 2 shows the workflow described in the study of Lasecki et al. (2013), however adjusting it to the workflow model proposed in our study.

The project presented in Lasecki et al. (2013) proposed a chat system that seeks to answer questions by crowdsourcing means. In this workflow, the workers initially completed a training stage; then the crowd executed the first crowdsourcing task, which consisted of proposing an answer to a question made in the chat system; the second crowdsourcing task consists of evaluating and voting the suggestions made by other workers in the first task. In the end, the majority decision activity selected the proposal with most votes to answer the question made in the chat system.

Regarding the creation of a crowdsourcing workflow, Kulkarni et al. (2012) presented an alternative method to divide a complex problem into small tasks, thus generating a crowdsourcing workflow. Retelny et al. (2017) suggest that if a complex problem can’t be adapted to the static crowdsourcing workflow approach, it will remain as a fundamental limitation of workflow-based crowdsourcing infrastructures. However, Fig. 1 shows the Object for Solution, which is an alternative crowdsourcing task type capable of solving complex problems, as presented in Section 4.3.

It’s worth pointing out that the second task of the workflow presented above was initially considered as a quality management activity, named in the Section 4.6 as subtasks. This second task still manages the quality of the first task results, but after defining the crowdsourcing workflow, we decided that the subtasks are better represented as a crowdsourcing task of object evaluation type.

5.3. About the crowdsourcing platform

The crowdsourcer must know the crowdsourcing platforms that are available to execute his project. These platforms, the tendencies of usage, and their limitations were presented in Section 4.5 of this study.

We also defined that the crowdsourcing platform must execute two primary functions: (i) Worker management and (ii) task management. We describe these features below:

1. Worker management: Function responsible for recruiting workers according to the criteria defined in a project, realizing filterings and forwarding the selected workers to the task management function. This feature is also responsible for distributing the rewards to the workers when these are extrinsic.
2. Task management: Function responsible for distributing crowdsourcing task to the workers and collect the task results.

It’s noteworthy that a single platform can execute these two features, as frequently done in projects that use the Mturk platform. However, distinct platforms can perform these functions, as approached by the CrowdFlower platform, which executes the task management while outsources the worker management to other general-purpose platforms, including the Mturk platform. Projects that use distinct platforms to run these two features are presented in the following studies: (McAllister Byun et al., 2015; Nguyen-Dinh et al., 2013).

6. Conclusion

This paper presented a systematic review of indexed papers, which aims to collect information and define characteristics about crowdsourcing projects and how their results are reported in the literature. Evaluating 76 crowdsourcing projects found in 72 articles, allowed us to broaden our view of the crowdsourcing field of study.

The tendencies and definitions revealed aim to help researchers and crowdsourcers to understand the state-of-the-art of crowdsourcing. These contributions are beneficial as the main concerns when designing further crowdsourcing projects.

Regarding all projects reviewed, we also observed that the concept of a crowdsourcing workflow in a project is not widespread. Thus, we found that the reviewed papers described and modeled their crowdsourcing projects and workflows in the way that they considered feasible. Therefore, this lack of standardization aroused our interest, as a future work, to update the definition of crowdsourcing projects and create a conceptual model that leverages the concept of a crowdsourcing workflow.
As another future work, we also wish to evaluate the quality management in crowdsourcing projects better, aiming to build a guideline of how the quality mechanisms cited should be used in each crowdsourcing project to get reliable results.

The limitations of our review are similar to the limitations of other systematic reviews as well. By considering indexed papers, some important material that could contain valid crowdsourcing projects was not included in this review, such as dissertations, white papers, short papers, and papers that we did not have access. Also, crowdsourcing projects with commercial purposes can have different configurations from crowdsourcing projects with academic purposes, thus, our point-of-view in this review is limited to crowdsourcing projects published in the literature.

Acknowledgments

The first author acknowledges FAPES (Process: 74010158) for the financial grant. Celso A. Saibel Santos is currently funded by CNPq (# 312148/2014-3) and FAPES (# 67927378/2015). Our gratitude goes to Ricardo de Almeida Falbo, who made important contributions to define the data gathering protocols used in our Review Process. Special thanks to Edmilson Geraldo Balbino Silva, who reviewed our English writing.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.ipm.2018.03.006

References

Dumitrache, A., Aroyo, L., Welty, C., Sips, R.-J., & Levas, A. (2013). “Dr. detective”: Combining gamification techniques and crowdsourcing to create a gold standard in medical text. In M. Acosta, L. Aroyo, A. Bernstein, J. Lehmann, N. Noy, & E. Simperl (Eds.), 1st international workshop on crowdsourcing the semantic web, CrowdSsem 2013 1030. 1st international workshop on crowdsourcing the semantic web, CrowdSsem 2013 (pp. 16–31). VU University Amsterdam, Netherlands: CEUR-WS.
