The Four Pillars of Crowdsourcing: a Reference Model

Mahmood Hosseini, Keith Phalp, Jacqui Taylor, and Raian Ali
Faculty of Science and Technology
Bournemouth University, UK
{mhosseini, kphalp, jtaylor, rali}@bournemouth.ac.uk

Abstract—**Crowdsourcing is an emerging business model where tasks are accomplished by the general public; the crowd. Crowdsourcing has been used in a variety of disciplines, including information systems development, marketing and operationalization. It has been shown to be a successful model in recommendation systems, multimedia design and evaluation, database design, and search engine evaluation. Despite the increasing academic and industrial interest in crowdsourcing, there is still a high degree of diversity in the interpretation and the application of the concept. This paper analyses the literature and deduces a taxonomy of crowdsourcing. The taxonomy is meant to represent the different configurations of crowdsourcing in its main four pillars: the crowdsourcer, the crowd, the crowdsourced task and the crowdsourcing platform. Our outcome will help researchers and developers as a reference model to concretely and precisely state their particular interpretation and configuration of crowdsourcing.**

Keywords—*crowdsourcing; crowdsourcing taxonomy; crowdsourcing configuration*

I. INTRODUCTION

Crowdsourcing is a new business model, which enables business owners to rely on the power of crowd to get jobs done. The participants could work independently or collaboratively to accomplish the task and they could be given incentives, tangible or intangible as a reward. Crowdsourcing is an alternative to outsourcing [1, 2, 3], which typically requires contracting expert suppliers and service providers [4, 5, 6]. Crowdsourcing is deemed particularly useful for tasks requiring a large number of viewpoints and problem solvers [7, 8, 9] and, at the same time, easy to perceive by the general public [10, 11, 12].

Though it can still be done in person [13, 14, 15], through meetings with the public and using audience response systems such as Turning Point, crowdsourcing usually takes place online. This is within the remit of the concept where tasks are often required to be completed quickly with little prior arrangement with participants. Online crowdsourcing also allows access to a larger and more diverse audience. That is to say, an online crowdsourcing platform is an information system which is not traditional, in the sense that it facilitates the work and acts as intermediate between stakeholders (the crowd and the crowdsourcer) who could be unknown to each other and are not adhering to the same organizational rules and roles for accomplishing tasks which are also not known a priori [16, 17, 18]. This makes the development of such

platforms challenging. Examples of commercial platforms for crowdsourcing include Amazon Mechanical Turk [19, 20], and Threadless [21, 22].

Crowdsourcing has been used in multiple domains in order to harness the power and wisdom of the crowd. Perhaps the most well-known example is Wikipedia. It has been also used in business and marketing [23, 24], medicine [25, 26], sociology [27, 28], and environmental sciences [29, 30]. Recently, crowdsourcing has been advocated for the domain of information systems analysis and design through the involvement of users in evaluating the software and contributing to update its models and also to evolve and maintain the software [31, 32, 33, 34]. Indeed, a pressing future research challenge is to decide how best to develop not only crowdsourcing platforms but also how best to utilize crowdsourcing, and its potential benefits, in the information systems development

Despite the variety of domains in which crowdsourcing has been proposed and the platforms that accommodate it, we still lack a shared understanding of the concept. A great deal of emphasis has been placed on the application of the concept for solving certain problems but, while doing that, there has been little attention to the ontological and conceptual foundations on how to engineer the entire process. The lack of consensus on the meaning and semantic of crowdsourcing has been recognized in [35]. The authors review the literature to provide a definition mainly in an accumulative way, which provides a definition that accommodates the main views of crowdsourcing.

Coming up with a common definition of crowdsourcing might be an unfeasible solution to bring together the variety of interpretations of crowdsourcing. There is a clear divergence in the literature on what should be considered core and optional and sometimes whether to consider it at all. For example, while the competence of the crowd in performing a crowdsourced task is considered core in [23, 36], it is not the case in [2, 37]. Another example relates to diversity, which is mentioned in [38, 39] but not mentioned in [40, 41]. This motivated us to think of a taxonomy of crowdsourcing which accommodates the diversity and commonality of the interpretation and perception of crowdsourcing in a variety of disciplines including computer science, sociology, and business.

In this paper, we review the literature on crowdsourcing and deduce a taxonomy for the concept in its four pillars: the crowd, the crowdsourcer, the task and the crowdsourcing platform. The taxonomy accommodates the diversity of views instead of trying to reach a consensus. We also elaborate on the meaning of the features and challenges they introduce when we engineer crowdsourcing platforms. This taxonomy will help as a reference model for developers and researchers to precisely define their specific interpretation of crowdsourcing, thus making its semantic clearer in the context of their work.

The paper is structured as follows. In Section 2, we discuss the importance of and need for a reference model for crowdsourcing, which is the essence of this study. In Section 3, we present our research method and the steps we took in order to construct a taxonomy model for crowdsourcing. In Section 4, we discuss the main four pillars of crowdsourcing and also present the constituents of these main pillars. In Section 5, we explore some cross-cutting dependencies among some of the building blocks of crowdsourcing pillars. Section 6 explains the research challenges in this field. We conclude the paper and present our future work in Section 7.

II. CROWDSOURCING: TAXONOMY NEEDED

A taxonomy is deemed critical because it brings order to a body of information or a collection of objects under study. A taxonomy very often enhances one's ability to understand complex concepts and ideas.

The function of a taxonomy is to classify or categorize a body of knowledge or collection of objects in a hierarchical manner. Each layer of the hierarchy "inherits" or possesses all attributes of the one immediately above - whatever those attributes might be. When one views a hierarchy from top to bottom, the matter becomes more particular and more specific the lower one goes in the hierarchy. Our taxonomy of crowdsourcing will serve as a reference model for future researchers and practitioners in this field of study.

The various definitions for crowdsourcing in the literature, plus the absence of a taxonomy for crowdsourcing, has led to some confusion both in the concept of crowdsourcing and in its usage. The definitions for crowdsourcing vary from one-line definitions [3, 42] to multi-sentence definitions [35, 43], with each definition listing a selection of features that may or may not be present in others. For example, we take three definitions of crowdsourcing from three different papers just to show how these definitions can vary, be partial to crowdsourcing, and sometimes even contradicting each other.

In our first definition, Erickson [44] defines crowdsourcing in the following way:

"By 'crowdsourcing' I mean: Tapping the perceptual, cognitive or enactive abilities of many people to achieve a well-defined result such as solving a problem, classifying a data set, or producing a decision."

In this definition, the crowd involved in crowdsourcing should have some perceptual, cognitive or enactive abilities, and it should consist of many people.

In our second definition, Li et al. [45] have used the following definition for crowdsourcing:

"Crowdsourcing has recently emerged as a powerful alternative. It outsources tasks to a distributed group of people (usually called workers) who might be inexperienced on these tasks."

In this definition the crowd is considered to be a distributed group of people. Furthermore, the crowd might be inexperienced in the task. These two definitions have almost nothing in common. Furthermore, they contradict each other when one paper sees competence (i.e. abilities) an essential feature of the crowd, while the other one observes it as an optional feature of the crowd. While both definitions can be true and worthwhile for the context they are applied in, researchers in the field of crowdsourcing may wonder which definition of crowdsourcing may best suit their purpose.

In our third definition, Faradani et al. [15] defined crowdsourcing in the following way:

"A crowdsourcing market is the place, usually an online website, where workers find and perform tasks often for a financial reward."

This last definition introduces two new features of crowdsourcing. In this definition, crowdsourcing is an online activity, and the crowd is often rewarded financially for performing such an activity. These features were not present in the first two definitions, while none of the features mentioned in the first two definitions are present in the third definition. Such heterogeneity in defining crowdsourcing was our first motivation in conducting this study.

Our second motivation was the usefulness of introducing taxonomy in a field of study. Many studies in different domains have used taxonomies in order to clarify and classify a concept under study. In computer sciences, for example, a taxonomy of reverse engineering was presented in [46], a taxonomy of autonomous agents was studied in [47], and a taxonomy of web search was introduced in [48]. The generality of utilizing the notion of taxonomy to clarify and classify different concepts also motivated us to conduct this study for a taxonomy of crowdsourcing.

III. RESEARCH METHOD

In order to construct the four pillars or crowdsourcing, we collected papers from various domains of study in which the crowdsourcing was utilized, since crowdsourcing is a multidisciplinary concept. We found crowdsourcing to be utilized in many fields of study, including, but not limited to computer science, business and management, medicine, environmental sciences and sociology. Studying and analysing these papers, we identified that crowdsourcing is comprised of four parts. These four parts, or four pillars, constitute the entire crowdsourcing operation and are as follows:

- The crowd: the crowd consists of the people who take part in a crowdsourcing activity.

- The crowdsourcer: the crowdsourcer is the entity (a person, a for-profit organization, a non-profit

organization, etc.) who seeks the power and wisdom of the crowd for a task at hand.

- The crowdsourcing task: the crowdsourcing task (simply called the task hereinafter) is the activity in which the crowd participates.

- The crowdsourcing platform: the crowdsourcing platform is the system (software or non-software) within which a crowdsourcing task is performed.

To fully understand the concept of crowdsourcing, we also performed a systematic mapping study and selected 113 papers. Our selection criteria allowed papers in which crowdsourcing was explicitly defined. These definitions helped us in feature extraction and subsection recognition. For more details about the selection process and the content analysis of the definitions please refer to **http://goo.gl/gsV5XB**.

After collecting and evaluating these papers, and identifying the four parts of crowdsourcing, we investigated these papers again in order to obtain more information on these four pillars. In particular, we looked at different definitions that these papers had suggested for crowdsourcing and tried to identify the features and subsections of these four crowdsourcing parts. We identified these features and subsections by performing a content analysis on the definitions. In our context analysis, it was necessary to be able to identify different words and phrases used in different definitions that pointed toward the same concept. For example, expressions like "disseminated competencies", "varying knowledge", "varying levels of expertise" and "different talents" were all translated into a feature called "expertise diversity."

Feature extraction and subsection abstraction were completed using a multi-stage method, i.e., we reviewed the definitions from the first to the last several times to ensure that no features and subsections are left unnoticed, and a second reviewer also confirmed the list of extracted features and subsections by going through the papers and definitions and approving the list of extracted features and subsections. When in doubt, the two reviewers discussed whether a particular feature or subsection should be extracted from a particular definition or not, and if they did not reach an agreement, a third reviewer was called to settle the dispute.

The next step in our research was to categorize and classify these features into well-known and well-accepted classifications. While sometimes it was possible to find these classifications in the literature, other times we had to create these classifications ourselves. To perform the latter, we got help from a second and third reviewer, and once more we went through a multi-stage revision process to ensure the quality of our classification.

IV. THE FOUR PILLARS OF CROWDSOURCING

As we mentioned earlier in this paper, we identified four main pillars of every crowdsourcing activity that were present in the current literature, plus we also identified the building blocks for these four pillars. In this section, we will analyse each pillar.

A. Pillar One: The Crowd

After reviewing the current literature on crowdsourcing, we recognized that the crowd of people who participate in a crowdsourcing activity have five distinct features. These features, and their definitions, are as follows:

- Diversity: diversity is the state or quality of being different or varied. In crowdsourcing, diversity means the recruitment of different people within the crowd to accomplish a task. Such diversity can be divided into four subcategories. Spatial diversity means recruiting the crowd with different backgrounds and from different locations. Gender diversity means recruiting the crowd without selection according to one's gender. Age diversity means recruiting the crowd without considering their age. Expertise diversity means recruiting the crowd of varied experience, expertise and competence.

- Unknown-ness: unknown-ness (or anonymity) is the condition or fact of being anonymous. In crowdsourcing, unknown-ness has one or two of these meanings. The first one is when the crowd participating in a task does not know the crowdsourcer. The second one is when the crowd participating in a task does not know other members of the crowd.

- Largeness: largeness means consisting of big numbers. Largeness also means being comprehensive. In crowdsourcing, largeness occurs when the crowd participating in a crowdsourcing activity is enough to fulfil and achieve the task. However, under certain circumstances, this largeness should not be so abundant to avoid overload, confusion, and unnecessary difficulty in management of the crowd.

- Undefined-ness: undefined-ness means not being determined and not having borders set. It means randomness. In crowdsourcing, crowd undefined-ness occurs when the crowd is chosen without imposed selection procedures to select a group of people, e.g., people with certain locations, certain abilities, and certain workplaces.

- Suitability: suitability means suiting a given purpose, occasion, condition, etc. In crowdsourcing, crowd suitability means the fit of the crowd for performing a crowdsourcing activity. Such fit could arise when the crowd has certain abilities to perform the task, when the crowd has the ability to collaborate with other members of the crowd, when the crowd volunteers to perform the crowdsourcing activity, and/or when the crowd is motivated in accomplishing a crowdsourcing task. Motivation can be intrinsic or extrinsic. Intrinsic motivations are a feature within the crowd, while extrinsic motivations come from the crowdsourcer, in the form of financial incentives or other forms of incentivizations. Intrinsic motivations are typically the stronger of the two, and can be in the form of mental satisfaction gained from performing a

crowdsourcing activity (e.g., the joy of designing a T-shirt cover in Threadless), self-esteem (e.g., knowing the fact that one can design an algorithm for a given problem), personal skill development (e.g., developing personal photography skills through competitions held in iStockPhoto), knowledge sharing through crowdsourcing (e.g., Wikipedia), and love of the community in which a crowdsourcing task is being performed (e.g., open-source communities).

Table 1 summarises the features of the crowd, and lists some of the papers from which these features were extracted. In this table and all following tables, we have tried to identify at most four references for every feature and sub-feature of crowdsourcing activity. However, some of these features and sub-features had fewer citing references. For example, we could find self-esteem as a feature of the crowd only in [35]. Furthermore, few of these features and sub-features could not be found explicitly in crowdsourcing definitions in the literature, which means the author(s) did not think of these features as a building block of crowdsourcing. We have illustrated these features and sub-features in the table by N/A. These features are sometimes counterparts of other features which were explicitly mentioned and hence we deduced them. Moreover, some features are meant for classification only, i.e. concept or aspect features. For example, suitability is an umbrella term for a collection of features and sub-features of the crowd, and was introduced to classify and categorize a number of features and sub-features. We have illustrated this kind of features in the table by C/L.

TABLE 1: LIST OF CROWD FEATURES

The Crowd	Mentioned In
1. Diversity	[30, 35, 38, 39]
1.1. Spatial Diversity	[49, 50, 51, 52]
1.2. Gender Diversity	N/A
1.3. Age Diversity	N/A
1.4. Expertise Diversity	[23, 36, 53, 54]
2. Unknown-ness	[2, 55, 56, 57]
2.1. Not Known to Crowdsourcer	[12, 16, 17, 18]
2.2. Not Known to Each Other	[16, 28, 58, 59]
3. Largeness	[3, 13, 23, 58]
3.1. Number Fulfils the Task	N/A
3.2. Number Not Abundant	[60]
4. Undefined-ness	[1, 4, 17, 57]
5. Suitability	C/L
5.1. Competence	[23, 29, 36, 43]
5.2. Collaboration	[1, 7, 57, 59]
5.3. Volunteering	[25, 35, 38, 54]
5.4. Motivation	[7, 61, 62, 63]
5.4.1. Mental Satisfaction	N/A
5.4.2. Self-Esteem	[35]
5.4.3. Personal Skill Development	[35, 64]
5.4.4. Knowledge Sharing	[29, 61]
5.4.5. Love of Community	[65]

B. Pillar Two: The Crowdsourcer

A crowdsourcer might be an individual, an institution, a non-profit organization, or a company that seeks completion of a task through the power of the crowd. After reviewing the current literature on crowdsourcing, we identified four distinct features of the crowdsourcer. These features, and their definitions, are as follows:

- Incentives provision: An incentive is something that stimulates one to take action, work harder, etc. It is a kind of stimulus or encouragement. In crowdsourcing, a crowdsourcer may provide incentives as a kind of extrinsic motivation for the crowd. Incentives can be classified in three categories. Probably the most prominent incentive in today's crowdsourcing market is financial incentives. Some crowdsourcing markets, such as Amazon Mechanical Turk, provide payments (sometimes also referred to as micro payments) for the completion of an advertised task. Another type of incentive is social incentives. Some crowds may take part in crowdsourcing activities in order to gain peer recognition [64] or public recognition [35, 61, 67]. The third form of incentives that a crowdsourcer may provide for the crowd is entertainment incentives. Here, the crowdsourcer may provide a form of enjoyment or fun in the crowdsourcing activity [61, 67] or may design a game around the crowdsourcing activity [68].

- Open Call: By open call, we mean an audition which is open to anyone who is willing to try out an act. In crowdsourcing, this means that the crowdsourcing activity is mainly open to the general public, and not only to a pre-selected few. An open call provides the opportunity for everyone in the crowd to participate in a crowdsourcing activity. This is synonymous to undefined-ness as a feature of the crowd, but it is from the perspective of the crowdsourcer, i.e. the crowdsourcer may provide an open call so that an undefined group of people participate in a crowdsourcing activity.

- Ethicality provision: An ethical act has to do with ethics or morality. It means conforming to moral standards, or to the standards of conduct of a given profession or group. In crowdsourcing, there are three acts that can be considered ethical during a crowdsourcing activity. The first one is for the crowdsourcer to provide an opt-out procedure. The crowdsourcer may provide an opt-out procedure for the crowd, so that the crowd has the right to stop the crowdsourcing activity at any time it wants. Such opting out may or may not affect the participant's condition. For example, the participant may not receive the intended incentive, or they may not be asked to participate in another crowdsourcing activity. Such opting out may or may not affect the crowdsourcer's condition, too. For example, the crowdsourcer may have to provide another open call to obtain the minimum required number of

participants in the crowdsourcing activity. Furthermore, this may affect the crowdsourcing time or resources. The second one is for the crowdsourcer to provide feedback to the crowd about the results of the crowdsourced activity, at least upon the crowd's request. This has been observed by some researchers to be an ethical approach [7]. The third one is for the crowdsourcer to ensure that the crowd will not be harmed during the crowdsourcing activity. This is especially important when the crowdsourcing activity is being performed in a real environment, not online.

- Privacy provision: Privacy means the quality or condition of being private. It also means one's private life or personal affair. In crowdsourcing, privacy means that the crowdsourcer should not disclose the crowd's personal and private information to other participants, other organizations and other entities. Privacy provision may also ensure the participating crowd's answers and crowdsourcing activities will not be revealed to others.

Table 2 summarises the features of the crowdsourcer, and lists some of the papers from which these features were extracted.

TABLE 2: LIST OF CROWDSOURCER FEATURES

The Crowdsourcer	Mentioned In
1. Incentives Provision	[35, 36, 39, 65]
1.1. Financial Incentives	[11, 16, 30, 50]
1.2. Social Incentives	[35, 61, 64, 67]
1.3. Entertainment Incentives	[61, 67, 68, 69]
2. Open Call	[1, 12, 57, 67]
3. Ethicality Provision	C/L
3.1. Opt-out Procedure	N/A
3.2. Feedback to Crowd	[7]
3.3. No Harm to Crowd	N/A
4. Privacy Provision	N/A

C. Pillar Three: The Crowdsourced Task

A crowdsourced task is an outsourced activity that is provided by the crowdsourcer and needs to be completed by the crowd. A crowdsourced task may take different forms. For example, it may be in the form of a problem, an innovation model, a data collection issue, or a fundraising scheme. The crowdsourced task usually needs the expertise, experience, ideas, knowledge, skills, technologies, or money of the crowd. After reviewing the current literature, we identified eight distinct features for the crowdsourced task. These features, and their definitions, are as follows:

- Traditional operation: By traditional operation, we mean the way the crowdsourced task would be done in an organization if it were not crowdsourced. Without the concept of crowdsourcing, the task would either be done by the employees of the organization, or it would be outsourced to another contractor, agent, company or organization to be completed.

- Outsourcing task: Outsourcing means transferring (certain manufacturing operations, administrative activities, etc.) to outside contractors, esp. so as to reduce one's operating costs. A crowdsourced task is usually a task that would otherwise be outsourced.

- Modularity: Modularity means designating units of standardized size, design, construction, etc. that can be arranged or fitted together in a variety of ways. A crowdsourced task may be an atomic task, but it is usually a more complex task that is broken down into micro tasks to be completed by the crowd.

- Complexity: Complexity is the condition or quality of being complex, not simple or consisting of two or more related parts. A crowdsourced task may be a simple task, or it may be a complex one. The difference between modularity and complexity is that some crowdsourced tasks may be complex, but they may be atomic tasks and not broken down to micro tasks.

- Solvability: Solvability is the capability to be solved. A crowdsourced task is usually a task that is simple enough for humans to be solved, but too complex for computers. Again, our perspective in this categorization differs from our previous categorizations, as a task may be a complex task (complexity: complex) which is not broken down into micro tasks (modularity: atomic) and is nonetheless simple for a human participant to complete (solvability: simple for humans).

- Automation characteristics: Automation, in manufacturing, means a system or method in which many or all of the processes of production, movement, and inspection of parts and materials are automatically performed or controlled by self-operating machinery, electronic devices, etc. A crowdsourced task is usually a task which is either difficult to automate (otherwise a computer would solve it instead of crowdsourcing it) or expensive to automate. The existence of such expense and/or difficulty of task automation opens the way for crowdsourcing. It is worth mentioning that this category focuses on automation, and not complexity or solvability of the task. However, if a task is too complex for computers, it also means that it is too difficult to automate.

- User-driven: A user-driven activity is one that is powered or controlled by users. A crowdsourced task is a user-driven task. Such user-driven activities fall into one of these three subcategories. Sometimes they are in the form of problem solving tasks, which means the crowd should provide a solution to a particular problem. Sometimes they can are innovation tasks, which means the crowd should generate ideas or create designs. Sometimes they are co-creation tasks, which means the crowd participates in a production process in order to create a product.

- Contribution type: When somebody contributes to something, it means that they have a share in bringing about (a result), or are partly responsible for something. A crowdsourced task may be in one of the two contribution forms. First, the contribution of the crowd in the crowdsourced task can be an individual contribution. This means that every participant of the crowd performs the crowdsourced task without helping or getting help from others. Second, the contribution of the crowd in the crowdsourced task can be a collaborative contribution. This means that different participants of the crowd perform the crowdsourced task by acting as a team, pooling resources and working in partnership.

Table 3 summarises the features of the crowdsourced task and lists some of the papers from which these features were extracted.

TABLE 3: LIST OF CROWDSOURCED TASK FEATURES

The Crowdsourced Task	Mentioned In
1. Traditional operation	[1, 4, 5, 6]
1.1. In-house	[1, 2, 18, 57]
1.2. Outsourced	[2, 41, 57, 70]
2. Outsourcing Task	[1, 3, 4, 5]
3. Modularity	[35]
3.1. Atomic Tasks	N/A
3.2. Divisible to Micro Tasks	[12, 40, 54, 65]
4. Complexity	[35, 71]
4.1. Simple Tasks	[11, 29, 37]
4.2. Complex Tasks	[10, 29, 56, 65]
5. Solvability	C/L
5.1. Simple for Humans	[10, 11, 12, 37]
5.2. Complex for Computers	[12, 56, 65, 72]
6. Automation Characteristics	C/L
6.1. Difficult to Automate	[50]
6.2. Expensive to Automate	[50]
7. User-driven	[73, 74]
7.1. Problem Solving	[44, 60, 65, 67]
7.2. Innovation	[23, 39, 62, 73]
7.3. Co-creation	[43, 64, 70, 73]
8. Contribution Type	C/L
8.1. Individual Contribution	[1, 36, 70, 75]
8.2. Collaborative Contribution	[1, 36]

D. Pillar Four: The Crowdsourcing Platform

The crowdsourcing platform is where the actual crowdsourcing task happens. While there are examples of real (offline or in-person) crowdsourcing platforms [14], the crowdsourcing platform is usually a website, or an online venue. After reviewing the current literature, we identified four distinct features for the crowdsourcing platform. These features, and their definitions, are as follows:

- Crowd-related interactions: Crowd-related interactions are interactions provided by the crowdsourcing platform between the crowd and the platform. These interactions include, but are not limited to:

  o Providing an enrolment mechanism for the crowd to enrol in the crowdsourcing platform.
  o Providing an authentication mechanism to authenticate the crowd.
  o Providing a declaration mechanism for the crowd to declare their skills and abilities.
  o Providing an assignment mechanism for assigning crowdsourced tasks to the crowd.
  o Providing an assistive mechanism to assist the crowd in different activities that happen in the crowdsourcing platform, e.g. helping the crowd to enrol, and helping the crowd to perform the crowdsourced task.
  o Providing a submission mechanism for the crowd to submit their results.
  o Providing a coordination mechanism to coordinate the crowd during crowdsourcing activities.
  o Providing a supervision mechanism to supervise the crowd during crowdsourcing activities.
  o Providing a feedback loop mechanism to give feedback to the crowd about their crowdsourcing activities.

- Crowdsourcer-related interactions: Crowdsourcer-related interactions are interactions provided by the crowdsourcing platform between the crowdsourcer and the platform. These interactions include, but are not limited to:

  o Providing an enrolment mechanism for the crowd to enrol in the crowdsourcing platform.
  o Providing an authentication mechanism to authenticate the crowd.
  o Providing a broadcast mechanism for the crowdsourcer to broadcast their crowdsourced task.
  o Providing an assistive mechanism to assist the crowdsourcer in different activities that happen in the crowdsourcing platform, e.g. helping the crowdsourcer to enrol, and helping the crowdsourcer to broadcast the crowdsourced task.
  o Providing a time negotiation mechanism for the crowdsourcer to negotiate the deadline or duration of the crowdsourced task with the crowd. This mechanism also allows the crowdsourcer to determine a deadline or a permitted duration without negotiation.
  o Providing a price negotiation mechanism for the crowdsourcer to negotiate the financial incentives or rewards of completing the crowdsourced task with the crowd. This mechanism also allows the crowdsourcer to determine a fixed price or a reward without negotiation.
  o Providing a verification mechanism for the crowdsourcer to verify the results which are obtained from the crowd.

- Providing a feedback loop mechanism to give feedback to the crowdsourcer about their crowdsourcing activities.
- Task-related facilities: Task-related facilities are facilities provided by the crowdsourcing platform about the crowdsourced task. These facilities include, but are not limited to:
  - Providing an aggregation mechanism to aggregate the results of a crowdsourced task. The outcome of such aggregation will be sent to the crowdsourcer for further verification, and may also be partially sent to the crowd as part of the feedback.
  - Hiding results obtained from one participant in the crowd from other participants.
  - Storing history of completed tasks, either for every task, for every crowdsourcer, for every participant, or a combination of those. Such history may be useful in deciding for future task assignments, or for preventing one participant from completing one certain crowdsourced task multiple times, etc.
  - Providing a threshold mechanism for the quality of the obtained results to ensure a minimum quality is met.
  - Providing a threshold mechanism for the quantity of the obtained results to ensure a minimum and/or maximum quantity is met.
- Platform-related facilities: Platform-related facilities are facilities provided by the crowdsourcing platform about the crowdsourcing platform itself. These facilities include, but are not limited to:
  - Providing an online environment which is inherent to online crowdsourcing. However, if crowdsourcing is performed in a real (offline or in-person) environment in a particular crowdsourcing activity, there should still be an environment in which crowdsourcing will take place.
  - Managing platform misuse, either by the crowd or by the crowdsourcer.
  - Providing an easy, feasible interface both for the crowd and the crowdsourcer to work.
  - Providing an attractive, appealing interface both for the crowd and the crowdsourcer to work.
  - Providing an interactive interface both for the crowd and the crowdsourcer to work.
  - Providing a payment mechanism to allow the crowdsourcer to pay a certain amount of money to the participants. If the reward is something other than money, the platform should also provide mechanisms for it. For example, if the participants should get a certain piece of software or mobile application for free in exchange for their task completion, the platform should provide a download mechanism for the participants.

Table 4 summarises the facilities of the crowdsourcing platform and lists some of the papers from which these facilities were extracted.

TABLE 4: LIST OF CROWDSOURCING PLATFORM FACILITIES

The Crowdsourcing Platform	Mentioned In
1. Crowd-related Interactions	C/L
1.1. Provide Enrolment	[30, 43, 76, 77]
1.2. Provide Authentication	[43]
1.3. Provide Skill Declaration	[30, 43]
1.4. Provide Task Assignment	[24, 30, 38, 78]
1.5. Provide Assistance	[60]
1.6. Provide Result Submission	[30, 36, 39, 79]
1.7. Coordinate Crowd	[59, 65, 80, 81]
1.8. Supervise Crowd	[78]
1.9. Provide Feedback Loops	[82]
2. Crowdsourcer-related Interactions	C/L
2.1. Provide Enrolment	[30, 43, 76, 77]
2.2. Provide Authentication	[43]
2.3. Provide Task Broadcast	[2, 16, 30, 79]
2.4. Provide Assistance	[60]
2.5. Provide Time Negotiation	[30]
2.6. Provide Price Negotiation	[30, 38]
2.7. Provide Result Verification	[36, 43]
2.8. Provide Feedback Loops	[82]
3. Task-related Facilities	C/L
3.1. Aggregate Results	[13, 65, 83]
3.2. Hide Results from Others	[58]
3.3. Store History of Completed Tasks	[43]
3.4. Provide Quality Threshold	[36, 84, 85]
3.5. Provide Quantity Threshold	[85]
4. Platform-related Facilities	C/L
4.1. Online Environment	[2, 3, 16, 57]
4.2. Manage Platform Misuse	[43]
4.3. Provide Ease of Use	[60]
4.4. Provide Attraction	[60, 85]
4.5. Provide Interaction	[58]
4.6. Provide Payment Mechanism	[30, 43, 86]

V. CROSSCUTTING DEPENDENCIES

In the previous section, we have analysed the different views of crowdsourcing and provided a taxonomy of the concepts which define each of its four pillars. In this section, we further enrich that taxonomy by discussing a set of inter-dependencies between those concepts to help the decision of crowdsourcing developers on configuring it. To express dependencies, we use the popular notions of *Require* and *Exclude* in Feature Model [87]. We also use two new relations *Support* and *Hinder*, which are the lighter versions of *Require* and *Exclude*. *Support* means that a feature empowers or facilitates another feature. *Hinder* means that a feature deters or discourages another feature. The reason why these two new relations are introduced is that unlike software features where interdependencies are crisp, crowdsourcing features are social and cognitive features and some interdependencies are qualitative.

- *Co-creation*, a task feature, **requires** *Collaboration*, a crowd feature. By definition, co-creation needs collaboration among participants.

- *Collaborative Contribution*, a task feature, **requires** *Collaboration*, a crowd feature. It also **requires** *Coordinate Crowd*, a platform facility.

- *Competence*, a crowd feature, **supports** *Complexity, Solvability, Problem Solving, Innovation*, and *Co-creation*, five task features. The reason is that a competent participant will better solve complex tasks and will be more efficient in user-driven crowdsourcing activities.

- *Complex for Computers*, a task feature, **requires** *Difficult to Automate*, another task feature.

- *Complex Tasks*, a task feature, **supports** *Divisible to Micro Tasks*, another task feature. The notion is that complex tasks are usually broken down into smaller tasks, which are called micro tasks.

- *Feedback to Crowd*, a crowdsourcer feature, **supports** *Motivation*, a crowd feature. Providing feedback on the results increases the crowd's trust in the system and motivates them to participate in future tasks.

- *Incentives Provision*, a crowdsourcer feature, **supports** *Largeness* and *Motivation*, two crowd features.

- *Largeness*, a crowd feature, **supports** *Diversity*, a crowd feature. It **hinders** *Coordinate Crowd*, a platform facility. Coordinating a large group of participants is difficult to achieve, especially for collaborative and complex tasks.

- *Not Known to Each Other*, a crowd feature, **excludes**, *Collaboration* another crowd feature. It also **excludes** *Collaborative Contribution*, a task feature. It **hinders** *Social Incentives*, a crowdsourcer feature, as the participants' identities will be hidden from each other.

- *Open Call*, a crowdsourcer feature, **supports** *Diversity, Largeness*, and *Undefined-ness*, three crowd features. It **requires** *Provide Task Broadcast*, a platform facility.

- *Opt-out provision*, a crowdsourcer feature, **supports** *Largeness, Volunteering*, and *Motivation*, three crowd features. The ability to opt-out without consequences would make the crowd more comfortable to participate.

- *Privacy Provision*, a crowdsourcer feature, **supports** *Not Known to Crowdsourcer* and *Not Known to Each Other*, two crowd features. It **requires** *Hide Results from Others*, a platform facility.

- *Provide Attraction*, a platform facility, **supports** *Largeness*, a crowd feature.

- *Provide Authentication*, a platform facility, **excludes** *Not Known to Crowdsourcer*, a crowd feature. It **supports** *Manage Platform Misuse*, another platform facility.

- *Provide Ease of Use*, a platform facility, **supports** *Largeness*, a crowd feature.

- *Provide Feedback Loop*, a platform facility, **supports** *Feedback to Crowd*, a crowdsourcer feature.

- *Provide Quality Threshold*, a platform facility, **requires** *Competence*, a crowd feature.

- *Provide Quantity Threshold*, a platform facility, **requires** *Largeness*, a crowd feature.

- *Provide Skill Declaration*, a platform facility, **supports** *Provide Task Assignment*, another platform facility. It **hinders** *Not Known to Crowdsourcer*, a crowd feature.

- *Simple for Humans*, a task feature, **supports** *Atomic Tasks* and *Simple Tasks*, two other task features.

- *Simple Tasks*, a task feature, **supports** *Atomic Tasks* and *Simple for Humans*, two other task features.

- *Suitability*, a crowd feature, **hinders** *Undefined-ness*, another crowd feature. The reason is that selecting a crowd with certain suitability features hinders the idea of undefined-ness, which is choosing the crowd on a random basis and without any imposed selection procedures.

- *Volunteering*, a crowd feature, **supports** *Motivation*, another crowd feature. The notion is that volunteering in an activity acts as an intrinsic motivation.

Our future work will develop a web-based platform for developers, which will assist them in configuring their crowdsourcing systems by giving recommendations and excluding inapplicable options based on the relations reported above. Another source of knowledge to assist this decision could be based on previous experiences of other developers and system analysts who could enrich the model of the four pillars with yet other metadata reflecting the good practice. In other ways, we would utilize crowdsourcing as a mechanism to build a knowledge base on how to develop such systems taking our taxonomy as a starting point.

VI. INFORMATION SYSTEMS AND CRWODSOURCING

While doing the review of the literature to deduce the taxonomy of crowdsourcing, we have also considered how crowdsourcing impacts information systems development and how it can be used to improve that process. We mainly observed that the systematic and principled crowdsourcing and the utilization of this mechanism for information systems and design is still under research.

We observed a lack of systematic approaches on engineering crowdsourcing. Considering crowdsourcing platforms as information systems we have not yet seen

development approaches which tackle the peculiarities of such systems and the challenges they introduce. There is a lack of holistic approaches and methodologies to build crowdsourcing platforms. The research has been mainly meant to address only separate facets of the development rather than providing methods, models and tools to facilitate the engineering itself. Examples of that include researches on motivations [88], pricing [9], coordination [89] and incentives [66]. Among other challenges, we also suggest other facets to accommodate in a comprehensive crowdsourcing analysis and design:

- **Metrics**: In the taxonomy we created in previous section, different concepts are still fuzzily defined. For example, we still need to define what largeness and diversity mean and how we measure whether the crowd is large and diverse enough. We speculate that this would be inherently relevant to the task being crowdsourced and the level of quality required and also the nature of the crowd recruited. Still, we have no metrics for that we still lack heuristics and patterns which could guide future development. Similar observations could be given to other dimensions such as motivation and social incentives.

- **Diversity**: Crowdsourcing platforms are inherently designed to be used by the crowd who could only partially be predicted, e.g. by defining what skills the participants should have. A similar observation could be made about the diversity of the crowdsourcers and the tasks. There is always a trade-off between accommodating diversity and achieving high quality. To maximize both, we may think of flexible and adaptive crowdsourcing platforms which could cater for such diversity and maintain quality. Research in adaptive systems [90] and perhaps Dynamic Software Product Lines [91] could be promising solutions here.

- **Recruitment**: Another trade-off relates to the openness to the crowd and ensuring that this openness will not reduce the quality of obtained knowledge. This has two sides. The first relates to the suitability of the crowd and how to define suitability in a crisp way, and the second relates to devising procedures to ensure what participants declare is true in terms of skills and applying their mind truthfully on the task being crowdsourced. Research in social network analysis, e.g. [92], could be helpful to tackle this first-class concern in crowdsourcing.

- **Interaction and aggregation**: The regulation of interaction and the aggregation of knowledge from the crowd is a known problem, especially when the outcome is a sort of collective decision or collective intelligence. This is a noted problem in the famous notion of Wisdon of Crowds [93]. While this has been researched intensively in the community of semantic web [94], datamining [95], as well as multi-agent systems [96], crowdsourcing intensively introduces the human factor to the problem space. This mainly means that interaction and aggregation can not be seen separately from users characteristics, motivation, incentives, and other concepts depicted earlier in our taxonomy.

Our other observation relates to the limited use of crowdsourcing in the development of information systems, though the potential benefits are still being recognized, given that the papers are published starting mainly in the last few years. In the following, we list a set of areas in which the power of crowdsourcing was utilized:

- **Adaptive systems engineering**: Software adaptation is meant to empower software ability to cater for the diversity of environment in which it operates. Uncertainty is inherent here and the reliance on software engineers to fully specify the system is somehow unrealistic. Crowdsourcing could be used to enrich the basic knowledge provided by software engineers and help the system obtain multiple stakeholders' perception of the system operations in the different context of use, so that it assists the system to optimize its adaptation. Ali et al. [31, 32] propose the notions of Social Adaptation and Social Sensing for acquiring the users perception on the role of the system in achieving their requirements and its quality. They propose to utilize that to make adaptation decisions.

- **User feedback for requirements knowledge**: Users feedback on software could help developers to better understand the requirements of the next release of the system. This feedback could be explicit, e.g. via forums, or implicit, e.g. through monitoring their patterns of use of the software. Pagano and Maalej [34] propose the effects of user feedback on software and requirements engineering teams. They signify the importance of user feedback content on the number of downloads a mobile phone application gets.

- **Stakeholders discovery**: In complex and dynamic systems, it is hard to identify the set of stakeholders and their roles and expertise and also their requirements. Crowdsourcing here would help identifying a comprehensive set of stakholders from an initial set of stakeholders specified by the analysts. Lim et al. [97] propose that the identification of stakeholders relevant to the system is not strightforward and propose a participatory approach to that. The work considers the set of stakeholders as a social network. The analysts could know only few members who will then recommend more members to the analysts and so on.

- **Testing**: Testing requires the generation of test cases and spotting errors and faults in the system, which is often a costly task. While this is a traditional problem in information system development, it is amplified when we build systems by integrating different services and components from different sources, working in heterogeneous environments and contexts of use. These are common characteristics of nowadays systems. Bell et al. [98] propose a game-like environment HALO (Highly Addictive, sociaLly

Optimized) Software Engineering to utilize gamification as a mechanisms to gather software testing results from the students. That is, crowdsourcing could be also used for testing and also made enjoyable through gamification techniques.

- **Requirements elicitation**: In software paradigms like Cloud Computing and Mobile Apps, the users set is highly diverse and unpredictable. This means relying on an elite group of users to understand what functionality and quality attributes to meet in the software is limited and also costly. We could harness the power of the crowd to understand their requirements as part of the requirements elicitation stage. CrowdREquire [60] is an example where the concept of crowdsourcing is proposed for requirements elicitation. Another approach is proposed by Knauss [99] in which the author proposes to harness the power of the largest stakeholder group, including the end-users, for requirements elicitation in a dynamic context. Hosseini et al. [100] propose to systematically analyze the impact of crowdsourcing features on the quality of requirements elicited via crowdsourcing.

- **Validation**: Validation and users testing for implemented systems share the same difficulties as mentioned above for requirements elicitation. This is also true as users might not maintain the same opinion when time passes due to the emergence of competitive solutions and the use of software in contexts which were not thought of at the engineering stage. There is also some research on the use of crowdsourcing for empirical studies and validation. In [101], Sayeed et al. make use of a simple rule-based technique for domain adaptation and generate data to validate it. Then they use crowdsourcing techniques to show that this strategy generates reliable results on data not seen by the rule designers.

The amount of research in using crowdsourcing for information system development is still limited. Even in the existing studies, the use of crowdsourcing is still at the level of advocating the potential of the paradigm and the use of the general principles. We observed a lack of methods and systematic approaches on how to decide the settings of crowdsourcing, how to design task-specific crowdsourcing platforms, and how to recruit the right crowd for the tasks related to the engineering process. Research on these areas is still needed to exploit the potential of crowdsourcing and maximize the quality of tasks performed. The ad-hoc approaches could lead to harmful results instead of supporting the development process.

VII. CONCLUSIONS AND FUTURE WORK

Having undertaken an extensive search in the literature, we have been able to provide an initial taxonomy for crowdsourcing. We have adopted a receptive approach and tried to put the different views together instead of restricting the concept to one definition. Our taxonomy is expected to be of use to both researchers and practitioners in the area as a reference model to clarify and potentially use as a tool for configuring crowdsourcing platforms. Our short-term future work mainly consists of:

- **Extending and revising taxonomy**: Due to the vast increase in studes on crowdsourcing, our taxonomy may become eventually less comprehensive. In addition, it may be that with the rapid change in hardware platforms, mobile computing and applications development, other mechanisms for data gathering and data use could extend our conceptual framework. We plan to provide an open access Wiki-like participatory approach to maintain the up-to-date nature of the taxonomy. That is, we will crowdsource the maintenance of our taxonomy of crowdsourcing.

- **Capturing the crowd**: Further work is needed to ensure we have the correct mechanisms to describe crowdsourcing, to capture its nuances and to ensure we can use our understanding of its nature to provide systems that fully capture the power of the crowd.

- **Combining data**: In addition, as with other data-rich approaches, we must have robust mechanisms for using the data intelligently, for example to allow us to reason about and resolve conflicts, to note particular clusters and to aggregate the information in meaningful ways and to ensure quality.

- **Tools**: There are already crowdsourcing platforms, but these have relatively limited capabilities other than for gathering data. Clearly, there is a signficant scope for developers to produce enhanced tools.

REFERENCES

[1] J. Howe, "Crowdsourcing: a definition," [Online], 2006, available from: http://www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html [Accessed 24 February 2014].

[2] V. Ambati, S. Vogel, and J. G. Carbonell, "Towards task recommendation in micro-task markets," in *Human computation*, 2011.

[3] Y. Baba and H. Kashima, "Statistical quality estimation for general crowdsourcing tasks," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 554-562, ACM, 2013

[4] R. Das and M. Vukovic, "Emerging theories and models of human computation systems: a brief survey," in *Proceedings of the 2nd international workshop on Ubiquitous crowdsouring*, pp. 1-4, ACM, 2011.

[5] V. Della Mea, E. Maddalena, and S. Mizzaro, "Crowdsourcing to mobile users: a study of the role of platforms and tasks," in *DBCrowd 2013*, 14, 2013.

[6] K. Mao, Y. Yang, M. Li, and M. Harman, "Pricing crowdsourcing-based software development tasks," in *Proceedings of the 2013 international conference on Software engineering*, pp. 1205-1208, IEEE Press, 2013.

[7] D. MacLean, K. Yoshida, A. Edwards, L. Crossman, B. Clavijo, M. Clark, ... and D. G. Saunders, "Crowdsourcing genomic analyses of ash and ash dieback–power to the people," *GigaScience*, *2*(1), 2, 2013.

[8] J. A. Redi, T. Hoßfeld, P. Korshunov, F. Mazza, I. Povoa, and C. Keimel, "Crowdsourcing-based multimedia subjective evaluations: a case study on image recognizability and aesthetic appeal," in *Proceedings of the 2nd ACM international workshop on Crowdsourcing for multimedia*, pp. 29-34, ACM, 2013.

[9] Y. Singer and M. Mittal, "Pricing mechanisms for crowdsourcing markets," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 1157-1166. International World Wide Web Conferences Steering Committee, 2013.

[10] J. M. Mortensen, M. A. Musen, and N. F. Noy, "Crowdsourcing the verification of relationships in biomedical ontologies," in *AMIA Annual Symposium*, 2013.

[11] J. Ross, L. Irani, M. Silberman, A. Zaldivar, and B. Tomlinson, "Who are the crowdworkers?: shifting demographics in mechanical turk," in *CHI'10 extended abstracts on Human factors in computing systems*, pp. 2863-2872, ACM, 2010.

[12] D. E. Difallah, G. Demartini, and P. Cudré-Mauroux, "Mechanical cheat: spamming schemes and adversarial techniques on crowdsourcing platforms," in *CrowdSearch*, pp. 26-30, 2012.

[13] J. Lehman and R. Miikkulainen, "Leveraging human computation markets for interactive evolution," in *International conference on Machine learning 2013 workshop: Machine learning meets crowdsourcing*, 2013.

[14] J. Howe, "The rise of crowdsourcing," [Online], 2006, available from: http://www.wired.com/wired/archive/14.06/crowds.html [Accessed 24 February, 2014].

[15] S. Faradani, B. Hartmann, and P. G. Ipeirotis, "What's the right price? pricing tasks for finishing on time," in *Human computation*, 2011.

[16] B. B. Bederson and A. J. Quinn, "Web workers unite! addressing challenges of online laborers," in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, pp. 97-106, ACM, 2011.

[17] M. Hirth, T. Hoßfeld, and P. Tran-Gia, "Anatomy of a crowdsourcing platform-using the example of microworkers.com," in *Fifth international conference on Innovative mobile and internet services in ubiquitous computing*, IMIS 2011, pp. 322-329, IEEE, 2011.

[18] D. R. Saunders, P. J. Bex, and R. L. Woods, "Crowdsourcing a normative natural language dataset: a comparison of amazon mechanical turk and in-lab data collection," *Journal of medical Internet research*, 15(5), 2013.

[19] M. Buhrmester, T. Kwang, and S. D. Gosling, "Amazon's mechanical turk a new source of inexpensive, yet high-quality, data?," *Perspectives on Psychological Science*, 6(1), pp. 3-5, 2011.

[20] A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with mechanical turk," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 453-456, ACM, 2008.

[21] D. C. Brabham, "Moving the crowd at Threadless: motivations for participation in a crowdsourcing application," *Information, Communication & Society*, 13(8), pp. 1122-1145, 2010.

[22] J. Wu, C. Damminga, K. K. Johnson, and H. Y. Kim, "Content analysis of online co-design community interactions: a case study of crowd-sourced Threadless," *Journal of Global Academy of Marketing*, 20(4), pp. 334-342, 2010.

[23] V. Chanal and M. L. Caron-Fasan, "How to invent a new business model based on crowdsourcing: the Crowdspirit® case," in *Actes de la Conférence Internationale de Management Stratégique, Nice*, pp. 1-27, 2008.

[24] P. Whitla, "Crowdsourcing and its application in marketing activities," *Contemporary Management Research*, 5(1), 2009.

[25] A. Foncubierta Rodríguez and H. Müller, "Ground truth generation in medical imaging: a crowdsourcing-based iterative approach," in *Proceedings of the ACM multimedia 2012 workshop on Crowdsourcing for multimedia*, pp. 9-14, ACM, 2012.

[26] B. Yu, M. Willis, P. Sun, and J. Wang, "Crowdsourcing participatory evaluation of medical pictograms using Amazon mechanical turk," *Journal of medical Internet research*, 15(6), e108, 2013.

[27] J. Heinzelman, R. Brown, and P. Meier, "Mobile technology, crowdsourcing and peace mapping: new theory and applications for conflict management," in *Mobile Technologies for Conflict Management*, pp. 39-53, Springer Netherlands, 2011.

[28] M. N. Wexler, "Reconfiguring the sociology of the crowd: exploring crowdsourcing," *International Journal of Sociology and Social Policy*, 31(1/2), pp. 6-20, 2011.

[29] H. Gao, G. Barbier, and R. Goolsby, "Harnessing the crowdsourcing power of social media for disaster relief," *Intelligent Systems,* IEEE, 26(3), pp. 10-14, 2011.

[30] P. Fraternali, A. Castelletti, R. Soncini-Sessa, C. Vaca Ruiz, and A. E. Rizzoli, "Putting humans in the loop: Social computing for Water Resources Management," *Environmental Modelling & Software*, 37, pp. 68-77, 2012.

[31] R. Ali, C. Solis, I. Omoronyia, M. Salehie, and B. Nuseibeh, "Social adaptation: when software gives users a voice," in *ENASE'12*, 2012.

[32] R. Ali, C. Solis, M. Salehie, I. Omoronyia, B. Nuseibeh, and W. Maalej, "Social sensing: when users become monitors," in *ESEC/FSE'11*, 2011.

[33] D. Pagano and B. Bruegge, "User involvement in software evolution practice: a case study," in *Proceedings of the 2013 international conference on Software engineering*, pp. 953-962, IEEE Press, 2013.

[34] D. Pagano and W. Maalej, "User feedback in the appstore: an empirical study," in *21st IEEE international requirements engineering conference*, pp. 125-134, IEEE, 2013.

[35] E. Estellés-Arolas and F. González-Ladrón-de-Guevara, "Towards an integrated crowdsourcing definition," *Journal of Information Science*, 38(2), pp. 189-200, 2012.

[36] Y. Zhao and Q. Zhu, "Evaluation on crowdsourcing research: current status and future direction," *Information Systems Frontiers*, pp. 1-18, 2012.

[37] B. Mellebeek, F. Benavent, J. Grivolla, J. Codina, M. R. Costa-Jussa, and R. Banchs, "Opinion mining of Spanish customer comments with non-expert annotations on mechanical turk," in *Proceedings of the NAACL HLT 2010 workshop on Creating speech and language data with Amazon's mechanical turk*, pp. 114-121, Association for Computational Linguistics, 2010.

[38] C. J. Ho, S. Jabbari, and J. W. Vaughan, "Adaptive task assignment for crowdsourced classification," in *Proceedings of the 30th international conference on Machine learning*, ICML-13, pp. 534-542, 2013.

[39] V. M. Ribiere and F. D. D. Tuggle, "Fostering innovation with KM 2.0," *Vine*, 40(1), pp. 90-101, 2010.

[40] M. Larson, P. Cremonesi, A. Said, D. Tikk, Y. Shi, and A. Karatzoglou, "Activating the crowd: exploiting user-item reciprocity for recommendation," in *the first workshop on Crowdsourcing and human computation for recommender systems, ACM Conference Series on Recommender Systems, ACM RECSYS*, 2013.

[41] E. Schenk and C. Guittard, "Crowdsourcing: what can be outsourced to the crowd, and why?," in *Workshop on open source innovation*, Strasbourg, France, 2009.

[42] F. Bry, F. Kneissl, T. Krefeld, S. Luecke, and C. Wieser, "Crowdsourcing for a geographical and social mapping of Italian dialects," in *SoHuman 2: 2nd international workshop on Social media for crowdsourcing and human computation, 2013.

[43] M. Vukovic, "Crowdsourcing for enterprises," in *2009 World conference on Services-I*, pp. 686-692, IEEE, 2009.

[44] T. Erickson, "Some thoughts on a framework for crowdsourcing," in *CHI 2011: workshop on Crowdsourcing and human computation*, pp. 1-4, 2011.

[45] H. Li, B. Yu, and D. Zhou, "Error rate analysis of labeling by crowdsourcing," in *International conference on Machine learning 2013 workshop: Machine learning meets crowdsourcing*, 2013.

[46] E. J. Chikofsky and J. H. Cross, "Reverse engineering and design recovery: a taxonomy," *Software, IEEE*, 7(1), pp. 13-17, 1990.

[47] S. Franklin and A. Graesser, "Is it an agent, or just a program?: a taxonomy for autonomous agents," in *Intelligent agents III agent theories, architectures, and languages*, pp. 21-35, Springer Berlin Heidelberg, 1997.

[48] A. Broder, "A taxonomy of web search," in *ACM Sigir forum*, Vol. 36, No. 2, pp. 3-10, ACM, 2002.

[49] D. A. Grier, "Not for all markets," in *Computer*, Vol. 44, No. 5, pp. 6-8, 2011.

[50] C. Heipke, "Crowdsourcing geospatial data," *ISPRS Journal of Photogrammetry and Remote Sensing, 65(6)*, pp. 550-557, 2010.

[51] P. Heymann and H. Garcia-Molina, "Turkalytics: analytics for human computation," in *Proceedings of the 20th international conference on World wide web*, pp. 477-486, ACM, 2011.

[52] R. Kern, H. Thies, C. Zirpins, and G. Satzger, "Dynamic and goal-based quality management for human-based electronic services," *International Journal of Cooperative Information Systems, 21*(01), pp. 3-29, 2012.

[53] K. T. Stolee and S. Elbaum, "Exploring the use of crowdsourcing to support empirical studies in software engineering," in *Proceedings of the 2010 ACM-IEEE international symposium on Empirical software engineering and measurement*, p. 35, ACM, 2010.

[54] D. L. Hansen, P. J. Schone, D. Corey, M. Reid, and J. Gehring, "Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at Familysearch indexing," in *Proceedings of the 2013 conference on Computer supported cooperative work*, pp. 649-660, ACM, 2013.

[55] T. Hossfeld, M. Hirth, and P. Tran-Gia, "Modeling of crowdsourcing platforms and granularity of work organization in future internet,"

in *Proceedings of the 23rd International teletraffic congress*, pp. 142-149, ITCP, 2011.

[56] B. Li, D. S. Appling, S. Lee-Urban, and M. O. Riedl, "Learning sociocultural knowledge via crowdsourced examples," in *Proceedings of the 4th AAAI workshop on Human computation*, 2012.

[57] J. O'Neill, S. Roy, A. Grasso, and D. Martin, "Form digitization in BPO: from outsourcing to crowdsourcing?," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 197-206, ACM, 2013.

[58] O. Alonso, D. E. Rose, and B. Stewart, "Crowdsourcing for relevance evaluation," in *ACM SigIR Forum*, Vol. 42, No. 2, pp. 9-15. ACM, 2008.

[59] W. Wu, W. T. Tsai, and W. Li, "Creative software crowdsourcing: from components and algorithm development to project concept formations," *International Journal of Creative Computing*, 1(1), pp. 57-91, 2013.

[60] A. Adepetu, K. A. Ahmed, Y. Al Abd, A. Al Zaabi, and D. Svetinovic, "CrowdREquire: a requirements engineering crowdsourcing platform," in *AAAI Spring Symposium Series*, 2012.

[61] A. Bozzon, M. Brambilla, S. Ceri, and A. Mauri, "Reactive crowdsourcing," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 153-164, International World Wide Web Conferences Steering Committee, 2013.

[62] M. K. Poetz and M. Schreier, "The value of crowdsourcing: can users really compete with professionals in generating new product ideas?," *Journal of Product Innovation Management*, 29(2), 245-256, 2012.

[63] L. Schmidt, "Crowdsourcing for human subjects research," in *Proceedings of CrowdConf*, 2010.

[64] D. C. Brabham, "Moving the crowd at iStockphoto: the composition of the crowd and motivations for participation in a crowdsourcing application," *First Monday*, 13(6), 2008.

[65] J. Goncalves, D. Ferreira, S. Hosio, Y. Liu, J. Rogstadius, H. Kukka, and V. Kostakos, "Crowdsourcing on the spot: altruistic use of public displays, feasibility, performance, and behaviours," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pp. 753-762, ACM, 2013.

[66] A. Singla and A. Krause, "Truthful incentives in crowdsourcing tasks using regret minimization mechanisms," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 1167-1178, International World Wide Web Conferences Steering Committee, 2013.

[67] G. Kazai, "In search of quality in crowdsourcing for search engine evaluation," in *Advances in information retrieval*, pp. 165-176, Springer Berlin Heidelberg, 2011.

[68] J. Pujara, B. London, and L. Getoor, "Reducing label cost by combining feature labels and crowdsourcing," in *ICML workshop on Combining learning strategies to reduce label cost*, 2011.

[69] W. Mason and D. J. Watts, "Financial incentives and the performance of crowds,"in *ACM SigKDD explorations newsletter*, 11(2), 100-108, 2010.

[70] J. Brito, "Hack, mash & peer: crowdsourcing government transparency," 2007.

[71] G. Hsieh, "Understanding and designing for cultural differences on crowdsourcing marketplaces," in *CHI 2011 workshop on Crowdsourcing and human computation*, 2011.

[72] M. Treiber, D. Schall, S. Dustdar, and C. Scherling, "Tweetflows: flexible workflows with twitter," in *Proceedings of the 3rd international workshop on Principles of engineering service-oriented systems*, pp. 1-7, ACM, 2011.

[73] S. Djelassi and I. Decoopman, "Customers' participation in product development through crowdsourcing: issues and implications," *Industrial Marketing Management*, 2013.

[74] R. Lukyanenko and J. Parsons, "Conceptual modeling principles for crowdsourcing," in *Proceedings of the 1st international workshop on Multimodal crowd sensing*, pp. 3-6, ACM, 2012.

[75] F. Kleemann, G. G. Voß, and K. Rieder, "Un(der) paid innovators: the commercial utilization of consumer work through crowdsourcing," *Science, Technology & Innovation Studies*, 4(1), pp. 1-5, 2008.

[76] C. Lofi, J. Selke, and W. T. Balke, "Information extraction meets crowdsourcing: a promising couple," *Datenbank-Spektrum*, 12(2), pp. 109-120, 2012.

[77] M. Porta, B. House, L. Buckley, and A. Blitz, "Value 2.0: eight new rules for creating and capturing value from innovative technologies," *Strategy & Leadership*, 36(4), pp. 10-18, 2008.

[78] V. Crescenzi, P. Merialdo, and D. Qiu, "Wrapper generation supervised by a noisy crowd," in *DBCrowd 2013*, 8, 2013.

[79] R. Cavallo and J. Jain, "Winner-take-all crowdsourcing contests with stochastic production," in *Workshop on Social computing and user generated content, EC-12, Valencia, Spain*, 2012.

[80] O. Alonso, "Perspectives on infrastructure for crowdsourcing," in *Crowdsourcing for search and data mining*, CSDM 2011, 7, 2011.

[81] R. Morris, "Crowdsourcing workshop: the emergence of affective crowdsourcing," in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. ACM, 2011.

[82] D. Oleson, A. Sorokin, G. P. Laughlin, V. Hester, J. Le, and L. Biewald, "Programmatic gold: targeted and scalable quality assurance in crowdsourcing," in *Human computation*, 11, 2011.

[83] I. Abraham, O. Alonso, V. Kandylas, R. Patel, S. Shelford, A. Slivkins, and H. Wu, "Crowdsourcing gold-HIT creation at scale: challenges and adaptive exploration approaches," in *HCOMP 2013 workshop*, 2013.

[84] R. M. Araujo, "99designs: an analysis of creative competition in crowdsourced design," in *First AAAI conference on Human computation and crowdsourcing*, 2013.

[85] D. C. Brabham, "Crowdsourcing as a model for problem solving an introduction and cases," in *Convergence: the international journal of Research into new media technologies*, 14(1), pp. 75-90, 2008.

[86] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin, "CrowdDB: answering queries with crowdsourcing," in *Proceedings of the 2011 ACM SIGMOD international conference on Management of data*, pp. 61-72, ACM, 2011.

[87] K. C. Kang, S. G. Cohen, J. A. Hess, W. E. Novak, and A. S. Peterson, "Feature-oriented domain analysis (FODA) feasibility study," Carnegie-Mellon University Pittsburgh, PA, Software Engineering Institute, 1990.

[88] J. Rogstadius, V. Kostakos, A. Kittur, B. Smus, J. Laredo, and M. Vukovic, "An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets," In *ICWSM*, 2011.

[89] A. Kittur, B. Smus, S. Khamkar, and R. E. Kraut, "Crowdforge: crowdsourcing complex work," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 43-52, ACM, 2011.

[90] B. H. Cheng, R. De Lemos, H. Giese, P. Inverardi, J. Magee, J. Andersson, ... and J. Whittle, "Software engineering for self-adaptive systems: a research roadmap," in *Software engineering for self-adaptive systems*, pp. 1-26, Springer Berlin Heidelberg, 2009.

[91] S. Hallsteinsen, M. Hinchey, S. Park, and K. Schmid, "Dynamic software product lines," *Computer*, 41(4), pp. 93-95, 2008.

[92] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proceedings of the 26th annual computer security applications conference*, pp. 1-9, ACM, 2010.

[93] J. Surowiecki, "The wisdom of crowds," Random House LLC, 2005.

[94] T. Gruber, "Collective knowledge systems: where the social web meets the semantic web," *Web semantics: science, services and agents on the World Wide Web*, 6(1), pp. 4-13, 2008.

[95] O. Z. Maimon and L. Rokach, "Eds.," *Data mining and knowledge discovery handbook*, Vol. 1, New York, Springer, 2005.

[96] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," in *Proceedings of the IEEE*, 95(1), pp. 215-233, 2007.

[97] S. L. Lim, D. Quercia, and A. Finkelstein, "StakeNet: using social networks to analyse the stakeholders of large-scale software projects," in *Proceedings of the 32nd ACM/IEEE international conference on Software engineering*, Vol. 1, pp. 295-304, ACM, 2010.

[98] J. Bell, S. Sheth, and G. Kaiser, "Secret ninja testing with HALO software engineering," in *Proceedings of the 4th international workshop on Social software engineering*, pp. 43-47, ACM, 2011.

[99] A. Knauss, "On the usage of context for requirements elicitation: end-user involvement in IT ecosystems," in *Proceedings of the 2012 IEEE 20th international requirements engineering conference*, RE '12, IEEE Computer Society, Washington, DC, USA, pp. 345-348, 2012.

[100] M. Hosseini, K. Phalp, J. Taylor, and R. Ali, "Towards crowdsourcing for requirements engineering," in *20th International working conference on Requirements engineering: foundation for software quality - Empirical Track*, 2013.

[101] A. B. Sayeed, T. J. Meyer, H. C. Nguyen, O. Buzek, and A. Weinberg, "Crowdsourcing the evaluation of a domain-adapted named entity recognition system," in *Human language technologies: the 2010 annual conference of the North American chapter of the association for computational linguistics*, pp. 345-348, Association for Computational Linguistics, 2010.