Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes

David Geiger
University of Mannheim
geiger@uni-mannheim.de

Stefan Seedorf
University of Mannheim
seedorf@uni-mannheim.de

Thimo Schulze
University of Mannheim
schulze@wifo.uni-mannheim.de

Robert Nickerson
San Francisco State University
RNick@sfsu.edu

Martin Schader
University of Mannheim
martin.schader@uni-mannheim.de

ABSTRACT

Crowdsourcing is an umbrella term for a variety of approaches that tap into the potential of a large and open crowd of people. So far, there is no systematic understanding of the processes used to source and aggregate contributions from the crowd. In particular, crowdsourcing organizations striving to achieve a specific goal should be able to evaluate the mechanisms that impact these processes. Following a method of IS taxonomy development we propose a new taxonomic framework for crowdsourcing processes. In contrast to previous work, this classification scheme focuses exclusively on an organizational perspective and on the mechanisms available to these organizations. The resulting dimensions are preselection of contributors, accessibility of peer contributions, aggregation of contributions, and remuneration for contributions. By classifying the processes of 46 crowdsourcing examples, we identify 19 distinct process types. A subsequent cluster analysis shows general patterns among these types and indicates a link to certain applications of crowdsourcing.

Keywords

Crowdsourcing, co-creation, classification, taxonomy, crowd, crowd management.

INTRODUCTION

In the last decade, the World Wide Web has evolved into a powerful medium for active collaboration among people located around the world. Many successful examples exist of people coming together on the Web to combine their resources – whether it is knowledge, creativity, opinions, skills, etc. – including the world's largest knowledge base Wikipedia, the problem solving platform InnoCentive, and the emergency coordination in response to the Haiti earthquake on Ushahidi. These phenomena are commonly referred to as crowdsourcing, a term coined by Jeff Howe and Mark Robinson in an article in Wired magazine (Howe, 2006a). Crowdsourcing can be defined as taking a function that is traditionally performed by employees and instead outsourcing it to the crowd "in the form of an open call". The crowd is an "undefined (and generally large) network of people" (Howe, 2006b).

Although the notion of crowdsourcing was introduced as recently as 2006, the idea existed before. As Howe states, crowdsourcing is "an umbrella term for a highly varied group of approaches that share one obvious attribute in common: they all depend on some contribution from the crowd. But the nature of those contributions can differ tremendously" (Howe, 2009). He defines four basic categories of crowdsourcing applications: crowd wisdom or collective intelligence; crowd creation or user-generated content; crowd voting; and crowdfunding. Crowdsourcing is a complex phenomenon, however, and often involves a combination of these categories, which may sometimes be hard to distinguish. This paper deals with any form of digital crowdsourcing, i.e., crowdsourcing via the Web, which has the potential to reach a large number of people with various backgrounds. Many crowdsourcing approaches make use of this potential diversity by using an open call format – also described as "self-identification of contributors" (Howe, 2009) – which allows anyone who is interested and capable to participate.

Figure 1 illustrates a prototypical crowdsourcing approach. A crowdsourcing organization relies on a particular crowdsourcing process to achieve an explicit goal. Two examples for crowdsourcing organizations are Amazon, which relies on crowdsourcing to provide comprehensive reviews for its products, and Threadless, which relies on crowdsourcing to create marketable t-shirt designs. The respective goal is realized by a process of sourcing and aggregating one or several kinds of contributions from the crowd (the crowdsourcing process). In the first example, the crowd is asked to contribute

either by writing new reviews or by commenting and voting on existing ones. These contributions are aggregated in what we will call an integrative approach and, thus, in most cases provide a comprehensive review of a product. In the second example, the crowd is asked to create new t-shirt designs and vote on existing ones. The design contributions are aggregated in a selective way, i.e., depending on the number of votes that a design receives, it is considered for production and the "winning" contributors are paid a share in profits.

Any organization that aims to adopt crowdsourcing in an effective way is required to carefully consider the characteristics of the crowdsourcing process that will be used for their particular goal. Existing literature on crowdsourcing does not cover this issue sufficiently. The purpose of this paper is to propose a systematic scheme for classifying crowdsourcing processes and, thus, identify the relevant mechanisms that impact these processes. Since crowdsourcing is used for a variety of different applications (product design, idea generation, problem solving, etc.), this paper focuses on those mechanisms that are applicable to all forms of crowdsourcing processes. To this end, we analyze the processes of a range of existing crowdsourcing examples on the Web and develop a taxonomy of these crowdsourcing processes. The classification of objects, i.e., the identification of similarities and differences among them, advances the understanding and analysis of complex domains and is considered one of the major strengths of taxonomies (Bailey, 1994). A taxonomy of crowdsourcing processes contributes to a better understanding of this domain. Moreover, it provides a foundation for future development of frameworks and methods to support the management of crowdsourcing processes.

The remainder of this paper is structured as follows: The next section analyzes and compares previous classifications in crowdsourcing and related areas. In section three, we describe our methodology for developing a taxonomy of crowdsourcing processes. Section four presents the resulting classification scheme. In section five, we identify clusters among types of crowdsourcing processes that were classified according to this scheme. Finally, we discuss the applications of our taxonomic framework and give directions for future research.

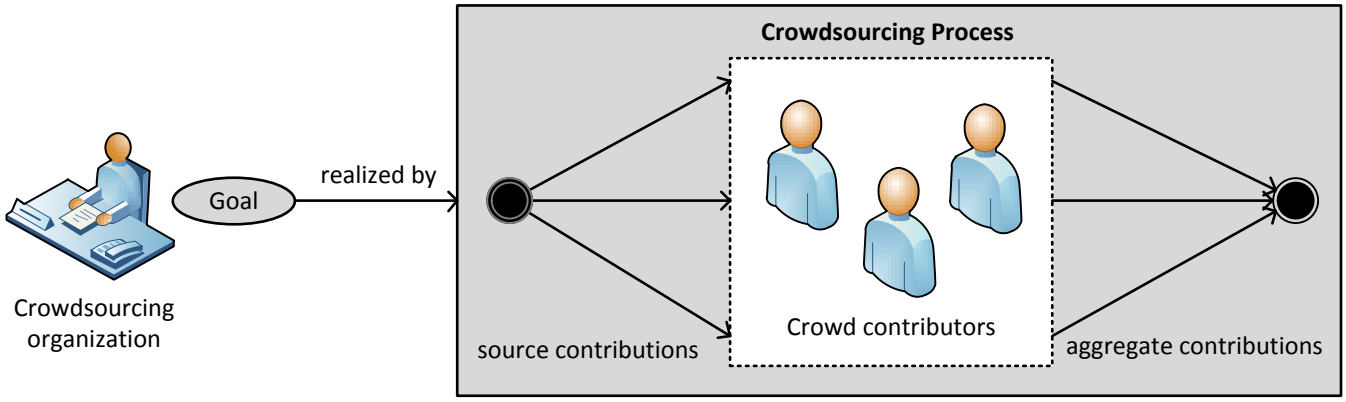

Figure 1. Prototypical crowdsourcing approach

RELATED WORK

Several classifications of crowdsourcing have been proposed in academic fields such as IS, economics, or management. To the best of our knowledge there is, however, no work that deals explicitly with the overarching processes of crowdsourcing approaches. Existing work often focuses on specific applications of crowdsourcing (e.g., open innovation or human computation) and does not consider crowdsourcing as a generic method. Most classifications do not deal with the processes of crowdsourcing but, for instance, with potential tasks (Kleemann, Voß, and Rieder, 2008), types of communities (Kozinets, Hemetsberger, and Schau, 2008; Whitla, 2009), or governance structures (Feller, Finnegan, Hayes, and O'Reilly, 2009). Other classifications do not take a unique perspective and relate to multiple concerns – mixing tasks, stakeholders, and processes – which partly elude direct control by a crowdsourcing organization. In the following section, we give an overview on such classification systems highlighting the process-relevant concerns. Table 1 summarizes these classifications.

In one of the most recent works, Doan, Ramakrishnan, and Halevy (2011) discuss crowdsourcing systems on the Web from a variety of perspectives. According to their definition, a crowdsourcing system "enlists a crowd of humans to help solve a problem defined by the system owners." In addition to classifying the characteristics of tasks and stakeholders in such systems, they also discuss several process-related aspects such as the explicit or implicit nature of collaboration or the combination and evaluation of inputs, i.e., crowd contributions.

Reference	Field	Motivation or purpose	Dimensions	Relates to
Doan et al. (2011)	IS, Computer Science	Global picture of crowdsourcing systems on the Web	Nature of collaboration	Process
			Type of target problem	Task
			How to recruit and retain users	Stakeholders
			What users can do	Task
			How to combine inputs	Process
			How to evaluate inputs	Process
			Degree of manual effort	Task
			Role of human users	Task/Stakeholders
			Standalone vs. piggyback	Process
Corney et al. (2009)	IS, Outsourcing	Foundation for identifying methodologies or analysis methods	Nature of the task	Task
			Nature of the crowd	Stakeholders/Task
			Nature of the payment	Process/Stakeholders
Schenk and Guittard (2011)	Management	Understanding crowdsourcing from a management science perspective	Integrative/selective nature of the process	Process
			Type of tasks	Task
Rouse (2010)	IS, Outsourcing	Clarifying the notion of crowdsourcing	Nature of the task / supplier capabilities	Task/Stakeholders
			Distribution of benefits	Stakeholders/Process
			Forms of motivation	Stakeholders/Process
Zwass (2010)	IS, Co-creation	Taxonomic framework as prerequisite for theory building in co-creation research	Autonomous vs. sponsored	Stakeholders/Task
			Performers	Stakeholders
			Motivation	Stakeholders
			Process governance	Process
			Task characteristics	Task
			Principal mode of product aggregation	Process
			Economic beneficiary	Stakeholders
Malone et al. (2010)	Collective Intelligence	Identifying the building blocks of collective intelligence approaches	What (goal)	Task
			Who (staffing)	Stakeholders
			Why (incentives)	Stakeholders
			How (structure/process)	Process
Piller et al. (2010)	Open Innovation	Analyzing strategies for customer participation in open innovation	Stage in innovation process	Task
			Degree of collaboration	Process
			Degrees of freedom	Task
Quinn and Bederson (2011)	Human Computation	A common understanding of human computation systems	Motivation	Stakeholders/Process
			Quality control	Process
			Aggregation	Process
			Human skill	Stakeholders/Task
			Process order	Process
			Task-request cardinality	Process

Table 1. Classifications of crowdsourcing approaches

Corney, Torres-Sanchez, Jagadeesan, and Regli (2009) find three possible dimensions for crowdsourcing. They first distinguish crowdsourcing tasks according to their nature: creation (e.g., design), evaluation (e.g., survey), and organization (e.g., tagging). Tasks are further categorized by the capabilities required. Some tasks can be solved by any individual, whereas others may require additional skills or expert knowledge. This dimension takes both a crowd and task perspective. The third dimension is concerned with the nature of the payment, thus dealing with the stakeholders as well as with the crowdsourcing process. In some cases the contribution is voluntary, in other cases the crowd (or parts of it) is rewarded with a fixed payment or a prize.

Schenk and Guittard (2011) define two dimensions in their typology: nature of the process and type of tasks. A crowdsourcing process may either pool complementary data (integrative nature) or give access to individual problem solving skills (selective nature). Crowdsourcing tasks are classified into routine, complex, or creative tasks. The authors then analyze the amount of remuneration for these tasks in several examples.

Rouse (2010) describes a preliminary taxonomy of crowdsourcing, the primary purpose of which is to clarify the meaning of the term. Crowdsourcing is mainly seen as a business solution and alternative form of outsourcing. The proposed taxonomy consists of three dimensions: supplier capabilities/nature of the task, distribution of benefits, and forms of motivation. The first dimension describes the complexity and skills involved in the task, while the second captures who benefits from crowdsourcing. The third subsumes the intrinsic and extrinsic motivational factors of the crowd. It is understood that motivation spans more than one motivational factor.

Crowdsourcing is related to the notion of co-creation. Zwass (2010) defines co-creation as the participation of consumers along with producers in the creation of value. He describes a "typology of co-created value", which is integrated into a "taxonomic framework of factors in co-creation." The typology distinguishes autonomous (e.g., Wikipedia) from sponsored co-creation (e.g., Threadless, Zazzle). The taxonomic framework then describes a number of dimensions that cover many different aspects including the stakeholders, the task characteristics, the co-creation process, and the co-created value.

Malone, Laubacher, and Dellarocas (2010) analyze 250 examples of collective intelligence and come up with a conceptual framework consisting of four building blocks ("genes"). They describe the "what", "who", "why", and "how" of collective intelligence approaches. The "what" gene distinguishes between a "create process" in which something new is generated and a "decide process" where alternatives are evaluated and selected. Individual crowd contributions can be independent of or dependent on each other. In the case of "creation processes" the former is called "collection" and the latter "collaboration". For the "decide process" there can be individual decisions or group decisions. The authors then describe different mechanisms for those genes.

Piller, Ihl, and Vossen (2010) present a typology of customer co-creation, which contributes to a better understanding of enterprise strategies for collaborative innovation. Their first dimension describes the stage in the innovation process that customers can participate in. The second dimension refers to the degree of collaboration between a firm and its customers and among the customers themselves. The third dimension describes the degrees of freedom that customers are given when working on a particular task. Based on these dimensions, eight types of customer co-creation are identified.

Quinn and Bederson (2011) introduce a taxonomy of human computation, which represents another specific area of application for crowdsourcing approaches. They classify human computation systems according to six dimensions: forms of motivation, quality control, aggregation of results, human skills, chronological order in which the stakeholders are involved, and task-request cardinality.

METHODOLOGY

Taxonomy development has been studied in a number of disciplines including biology (Eldredge and Cracraft, 1980; Sokal and Sneath, 1963), the social sciences (Bailey, 1994), and marketing, finance, and other areas of economics (Bock, Gaul, and Vichi, 1995-2011). Nickerson, Muntermann, Varshney, and Isaac (2009) and Nickerson, Muntermann, and Varshney (2010) examined taxonomy development in information systems. Their survey of 65 papers in different disciplines shows that there are different understandings of the term taxonomy depending on the literature. They define a taxonomy as a set of dimensions each consisting of a set of mutually exclusive and collectively exhaustive characteristics that describe how the objects under consideration differ. The term taxonomy can also refer to the actual result of classifying objects within such a taxonomic framework. Nickerson et al. propose a taxonomy development method that combines both empirical-to-deductive and deductive-to-empirical approaches to identify the dimensions (or variables) and corresponding characteristics (or variable domains). In their method, the dimensions and characteristics in the taxonomy follow from an overall or meta-characteristic, which derives from the users and purpose of the taxonomy, and serves as a basis for selecting characteristics in the sense that each characteristic is a logical consequence of the meta-characteristic.

The taxonomy development process starts with defining its users and purpose. This paper takes an organizational perspective on crowdsourcing. Thus, the intended user of the crowdsourcing taxonomy is the organization that wants to apply a crowdsourcing process to realize a certain goal and has to decide on the specifics of this process. From this perspective, it does not matter whether the decisions are taken in-house or are outsourced to an intermediary. A crowdsourcing organization may therefore realize a crowdsourcing process on its own or entrust an intermediary to do so.

The purpose of the taxonomy is to distinguish among crowdsourcing processes, i.e., the processes of sourcing contributions from the crowd and aggregating them in a particular way to achieve a predefined goal. The meta-characteristic that derives directly from this purpose is the mechanisms that impact the crowdsourcing process. We consider only aspects that can be directly influenced by the crowdsourcing organization. Furthermore, these mechanisms must apply to all kinds of crowdsourcing processes, whether they are used to aggregate the world's knowledge, to create t-shirt designs, or to generate new product ideas. As described later, we tested several commonly used dimensions from the crowdsourcing literature for compatibility with our meta-characteristic and integrated them where it seemed reasonable.

The objects that are classified in the taxonomy are Web-based crowdsourcing processes. We have identified a set of processes by analyzing a range of crowdsourcing examples from some of the existing literature (Corney et al., 2009; Howe, 2006a, 2009; Rouse, 2010; Schenk and Guittard, 2011; Zwass, 2010). These examples consist of organizations and platforms that use a crowdsourcing approach to achieve an explicit goal. For this paper we have analyzed the processes of 46 examples that are commonly accepted as crowdsourcing in the literature. This includes the most recognizable and often cited crowdsourcing examples across the reviewed literature as well as a randomly selected sample. The complete list can be found in Table 2 at the end of this paper. We are aware that the taxonomy may need to be extended or adapted, as more crowdsourcing examples and their processes are considered.

Most of the crowdsourcing processes in our list were derived from examples of specific organizations applying a crowdsourcing approach. Examples are TripAdvisor providing descriptions and evaluations of hotels, etc. through user reviews and ratings, or iStockphoto building up a stock of digital photographs provided by contributing individuals. Some crowdsourcing processes, however, may be realized using a generic crowdsourcing platform. A generic crowdsourcing platform realizes one or several kinds of repeatable and well-defined crowdsourcing processes that feature the same fixed characteristics. An example for a crowdsourcing platform is 99designs, which offers its customers a predefined process to crowdsource the design of logos, websites, etc. The crowdsourcing organization in this case is any organization that makes use of the particular platform. If a platform offers small variations in process characteristics, we consider these variations as different processes. The InnoCentive platform, for instance, where a crowdsourcing organization can have a particular problem solved, offers two possibilities regarding the crowd of solvers: asking anyone in the world or ask only their own employees. In order to reflect this basic difference, we analyze two separate processes here: that of the "InnoCentive Challenge Center" and that of "InnoCentive@Work".

As proposed by Nickerson et al. (2009), we followed an iterative approach to develop the taxonomy. In the first iteration, we analyzed the dimensions used in the existing crowdsourcing literature and took those related to our meta-characteristic. We added further dimensions that seemed to reflect important mechanisms with impact on the crowdsourcing process. We then used these dimensions to classify the processes in our list, based on our understanding of the latter. By classifying these processes, we identified further relevant differences between them, which needed to be reflected in the dimensions and their characteristics. Others proved to be irrelevant or redundant. Classification was performed independently by three of the authors in an effort to check inter-coder reliability. When the entire set of objects had been classified, we identified the discrepancies between the authors' classifications and analyzed their causes. In some cases, a mismatch was due to a misunderstanding and the description of the corresponding dimension was refined. In other cases, it demonstrated a problem with the dimensions themselves and led to the discovery of new or the refinement of existing ones. This entire development process was repeated until no additional or redundant dimensions could be identified.

CHARACTERISTICS OF CROWDSOURCING PROCESSES

By applying the above approach, we identified four dimensions that describe how crowdsourcing processes differ. Figure 2 depicts these dimensions and their temporal relevance within an implemented crowdsourcing process. First, the crowdsourcing organization must preselect the crowd of potential contributors to the process. Then the organization needs to decide how peer contributions will be disclosed as they are proposed. When the organization receives the results, it must aggregate the contributions. Finally the organization may need to compensate the contributors. This sequence is independent of the order in which a crowdsourcing organization decides on the respective characteristics when it plans the process. The exact shape of the decision process depends on the specific application, organization, etc.

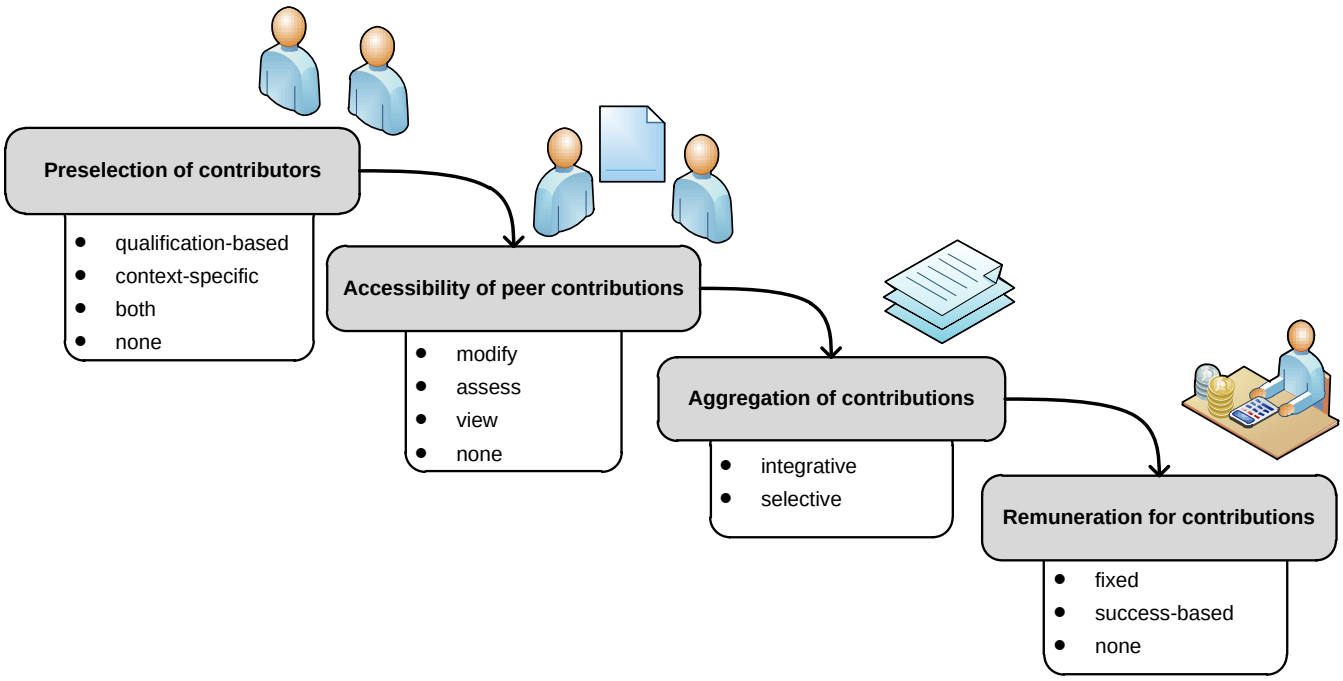

Figure 2. Characteristics of crowdsourcing processes

Preselection of contributors

The first dimension, *preselection of contributors*, is concerned with restrictions regarding the group of potential contributors. When a preselection is applied, a crowdsourcing organization places an open call to a restricted group of potential contributors who then decide if they want to contribute to the crowdsourcing process. Most crowdsourcing processes that we identified strive to benefit from as much diversity and scalability as possible and, thus, do not limit the contributors. Although *no* preselection is exercised in these cases, contributions and contributors may still be excluded ex post if they do not adhere to the respective rules and quality standards.

As a means of ensuring a minimum ex ante quality level of contributions, some processes require their contributors to demonstrate certain knowledge or skills before being allowed to regularly contribute. In our terminology, these processes apply a *qualification-based* preselection. Examples for contributor-level qualifications can be found on the 99designs logo store (contributors need to have won one design contest) and on iStockphoto (applicants need to upload sample pictures).

Another form of contributor preselection is due to *context-specific* reasons. Some organizations, for instance, restrict the crowd that is allowed to contribute to their own employees (e.g., InnoCentive@Work) or to their customers (e.g., e-Rewards). In the first case, this may be due to available implicit knowledge or privacy concerns. In the second case, organizations are only interested in their customers' opinions.

Finally, some processes apply *both* kinds (qualification-based and context-specific) of preselection. For instance, LiveOps, which keeps a pool of crowdsourced call center agents, requires potential contributors to meet different kinds of qualifications (product knowledge, reputation scores, etc.), and applies context-specific restrictions, e.g., on demographics (country, age group, etc.).

A similar distinction of the "performers" in a co-creation process can be found in Zwass (2010). He distinguishes between "the world", i.e., a completely open crowd, "prequalified individuals" or "skilled contributors", and "community members". This distinction, however, does not focus explicitly on those mechanisms that a crowdsourcing organization can directly influence. In the open innovation field, Feller, Finnegan, Hayes, and O'Reilly (2010) distinguish two general strategies for assembling a community of problem solvers: maximizing the size of the community, i.e., applying no preselection or pre-filtering the community by defining "a set of desirable skills and backgrounds".

Accessibility of peer contributions

The second dimension, *accessibility of peer contributions*, indicates to what extent contributors can access each other's contributions. The four characteristics of this dimension reflect the degree of access that a crowdsourcing process enables. They are, in order of increasing accessibility: none, view, assess, or modify. *None* means that contributors cannot see each other's contributions. Contributions are isolated from each other and there is no way contributors can reuse, complement, or otherwise react to another contribution. Crowdsourcing organizations may not want to allow visibility of peer contributions due to reasons such as privacy (e.g., customer surveys on e-Rewards), ensuring diversity (e.g., private contests on idea bounty and 99designs, or image tagging on Camclickr and the Google Image Labeler), or simply because it is not necessary for their purpose (e.g., LiveOps). In the case of private contests, contributions are only visible to the contest initiator, though they may be revealed after the contest is over.

On the lowest level of actual accessibility, the *view* characteristic means that all contributions are visible to any potential contributor. Two exemplary cases are public design contests (e.g., on 99designs or crowdspring) and customer reviews on Angie's List. Since the dimension represents the maximum degree of accessibility, there is no means for explicitly rating or commenting on other contributions.

If a crowdsourcing process uses a means for rating or commenting on other contributions, we characterize the accessibility of peer contributions as *assess*. Contributors of these processes can use explicit mechanisms to express their opinion on individual contributions. These mechanisms are often used to assess contributions in contests (e.g., on Atizo or InnoCentive@Work) as well as on all kinds of platforms around user-generated content "to sift through the flood of contributions that often follow crowdsourcing's open call" (Howe, 2009). Among the many examples are several kinds of digital stores (e.g., iStockphoto or the Android Market), social media sites (e.g., YouTube or the Yahoo! Contributor Network), and in general platforms that rely on user reviews and/or ratings (e.g., TripAdvisor or Delicious).

Finally, the highest level of accessibility is *modify*. Contributors in such processes can alter or even delete each other's contributions in order to correct, update, or otherwise improve them. In general, this is the case when contributors come together to build something in a highly collaborative way. Examples include Wikis, e.g., Wikipedia, and similar endeavors such as OpenStreetMap or the Emporis Community.

Aggregation of contributions

The third dimension, *aggregation of contributions*, describes how the crowd contributions within a crowdsourcing process are used by the crowdsourcing organization to achieve the desired outcome. This is based on what Schenk and Guittard (2011) call the "nature of the process." When a process is designed around *integrative* contributions, "the issue is to pool complementary input from the crowd." All contributions are reused for the final outcome unless they fail to meet certain quality requirements. Integrative crowdsourcing is mainly used to tap into the creative power (e.g., the Android market, iStockphoto, YouTube, and Wikipedia) or the collective opinion (e.g., Delicious, Digg, and the Google Image Labeler) of the crowd.

Selective crowdsourcing processes, by contrast, follow a more competitive approach to achieve their outcome. Individual contributions are compared to each other and the "best" one(s) is selected. According to Schenk and Guittard (2011), this approach allows a crowdsourcing organization "to choose an input from among a set of options that the crowd has provided." Selective crowdsourcing is mainly used in contests, e.g., 99designs, Atizo, InnoCentive, or the Netflix Prize. If the process involves a crowd assessment, the selection is sometimes based on the collective opinion. Threadless, for instance, claims that the overall community score that a submitted t-shirt design achieves within a seven-day period is used as a "gauge" to decide whether it goes into production. The crowdsourcing process used on the Dell IdeaStorm platform also relies to a considerable extent on the popularity that particular ideas receive among the community.

Remuneration for contributions

The fourth dimension, *remuneration for contributions*, determines how contributors are paid or otherwise compensated for their work. We distinguish between fixed, success-based, and no remuneration. *Fixed* remuneration means that all contributions that adhere to the respective terms and conditions generate a fixed payment regardless of their value to the final outcome. In the crowdsourcing processes we have identified so far, fixed remuneration is only applied to integrative cases (e.g., e-Rewards and LiveOps). *Success-based* remuneration means that contributions will be paid depending on their individual value to the crowdsourcing goal. Success may be obvious in selective forms of crowdsourcing, like a contest on 99designs or the InnoCentive Challenge Center, where only the winning contribution (in this case a design or a solution, respectively) is paid for. Success-based remuneration may also be applied in integrative cases like on iStockPhoto or the Yahoo! Contributor Network, where contributions are continuously paid based on the revenue they achieve. Besides fixed

and success-based payment, some crowdsourcing processes offer *no* remuneration at all and, thus, completely depend on other mechanisms to attract contributors.

Related dimensions in the academic literature are mainly concerned with the amount or the exact form of the payment (see "remuneration" in Schenk and Guittard (2011)) or the overall motivation of contributors (see "nature of the motivation to participate" in Rouse (2010) and "motivation" in Zwass (2010) and Quinn and Bederson (2009)). Motivation emerges as an ex post result of a particular crowdsourcing realization seen from a contributor's perspective. Most motivational factors, especially intrinsic ones such as passion, fun, community identification, or personal achievement, cannot be directly controlled by the crowdsourcing organization (Leimeister, Huber, Bretschneider, and Krcmar, 2009). In addition, motivational factors often overlap and are, thus, sometimes impossible to distinguish (Ryan and Deci, 2000). Certain mechanisms, however, when properly implemented, may influence the motivation of the crowd in an indirect way. An organization may provide incentives and create the right prerequisites for certain kinds of motivation to (potentially) emerge. Apart from remuneration, such mechanisms are mostly specific to different applications of crowdsourcing – e.g., idea competitions (Leimeister et al., 2009), user-generated content (Brabham, 2008; Schroer and Hertel, 2009), or design competitions (Brabham, 2010) – and therefore exceed the scope of this taxonomy.

The one dimension we have identified that has the closest similarity to our concept of remuneration, is the "nature of the payment" dimension applied by Corney et al. (2009). "Rewarded contribution at a flat rate" is equivalent to our characteristic of fixed remuneration. "Rewarded contribution with a bonus or prize" can roughly be compared to a success-based remuneration, but does not include the aspect of continuous profit sharing. "Voluntary contribution," finally, is equivalent to no payment, although it implies a motivational factor that we have tried to avoid for the above-mentioned reasons.

TYPES OF CROWDSOURCING PROCESSES

The proposed taxonomic framework describes 96 possible combinations of process characteristics. These combinations are called process types. Classification of the crowdsourcing processes used in 46 examples resulted in a list of 19 distinct types. In order to identify general patterns among these process types, we performed a cluster analysis on the processes of our examples. SPSS Statistic 19 was used to perform TwoStep Cluster Analysis with log-likelihood distance measure and the Schwarz's Bayesian cluster criterion. The number of clusters was five, which offers a satisfactory cluster quality (a silhouette measure of cohesion and separation) for our current sample of 46 processes. Although a larger sample set may result in a slightly different clustering, our current results provide interesting insights into the link between crowdsourcing processes and applications of crowdsourcing. Table 2 illustrates the resulting clusters along with the 19 process types and the crowdsourcing examples that apply the corresponding processes. The table is ordered by descending cluster size.

Taking a closer look at the five clusters, we were able to identify definitive characteristics and assign corresponding names. *Integrative sourcing without remuneration* is applied in many cases, e.g., for various forms of Wikis, user reviews, image tagging, or free user-generated content. *Selective sourcing without crowd assessment* comprises private (contributors do not see each other's contributions) and public design and innovation contests, in which one or a few winners are remunerated. *Selective sourcing with crowd assessment* refers to contests that allow fellow contributors and other people to publicly assess individual contributions. *Integrative sourcing with success-based remuneration* is mainly used on store platforms that sell user-generated content (e.g., software, photographs, and designs) on the basis of profit sharing. *Integrative sourcing with fixed remuneration*, finally, is often applied to transactional tasks or micro-tasks, varying in complexity and often restricting the crowd of potential contributors. Contributions are completely isolated from each other.

The predictor importance of the cluster analysis suggests that remuneration (1.0) and aggregation (0.58) have the most influence on cluster membership. Preselection (0.26) and accessibility (0.2) are less important in terms of distinguishing the clusters. This suggests that the two primary decisions that a crowdsourcing organization must make with respect to the crowdsourcing process are how the contributions are aggregated (integrative or selective) and how the contributors are remunerated. The other two dimensions, preselection and accessibility, give the crowdsourcing organization more detailed possibilities to adjust the process for specific applications.

CONCLUSION AND FUTURE WORK

The purpose of this paper is to propose a systematic classification of crowdsourcing processes and, thus, contribute to a better understanding of the mechanisms that determine these processes. Existing classifications mix different aspects of crowdsourcing approaches – not all of which can be directly influenced by a crowdsourcing organization – or focus on specific applications and, thus, lack general applicability. Based on empirical data and current literature, we have developed a taxonomic framework of crowdsourcing processes. We identified four dimensions that impact the process of sourcing and aggregating contributions from the crowd. By applying this classification scheme to the processes of a range of existing

crowdsourcing examples, 19 distinct process types were identified. Using a cluster analysis, we described five general patterns among these types and linked them to exemplary applications of crowdsourcing.

The taxonomy contributes to a more systematic understanding of the generic processes in crowdsourcing approaches. It can be useful for managers to get an overview of the basic choices when planning a crowdsourcing project. There are, however, a number of additional mechanisms, e.g. quality management, whose shape depends on the specific applications of crowdsourcing and which are not part of this framework. The classification of 46 crowdsourcing examples provided initial hints on which combinations of mechanisms are feasible or represent "best practices" for specific applications, e.g., idea generation or creative design. Further research is needed to study the relationships between applications and processes.

One direction for future research is to refine the taxonomy of crowdsourcing processes by classifying further instances. Only 19 of the 96 theoretically possible process types (for the current dimensions) have been identified so far. An extended analysis of the domain may lead to changes in the dimensions, the identification of further process types, or explain why certain combinations of process characteristics do not occur. An interesting approach for that effect may be to crowdsource the identification of crowdsourcing examples and the classification of their processes itself. Another direction is to explore the link between mechanisms of crowdsourcing processes and individual applications or categories of applications, e.g., as proposed by Howe (2009). Given the apparently similar applications within individual clusters in Table 2, this seems promising. Research in this direction may contribute towards answering the question what good crowdsourcing processes for a given application are.

REFERENCES

1. Bailey, K. D. (1994) Typologies and taxonomies: an introduction to classification techniques, Sage Publications, Inc.
2. Bock, H., Gaul, W., and Vichi, M. (1995-2011) Series "Studies in Classification, Data Analysis, and Knowledge Organization", Springer.
3. Brabham, D. C. (2008) Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application, *First Monday*, 13, 6, 1-22.
4. Brabham, D. C. (2010) Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application, *Information, Communication & Society*, 13, 8, 1122-1145.
5. Corney, J. R., Torres-Sanchez, C., Jagadeesan, A. P., and Regli, W. C. (2009) Outsourcing labour to the cloud, *International Journal of Innovation and Sustainable Development*, 4, 4, 294-313.
6. Doan, A., Ramakrishnan, R., and Halevy, A. Y. (2011) Crowdsourcing systems on the World-Wide Web, *Communications of the ACM*, 54, 4, 86.
7. Eldredge, N., and Cracraft, J. (1980) Phylogenetic Patterns and the Evolutionary Process, Columbia University Press.
8. Feller, J., Finnegan, P., Hayes, J., and O'Reilly, P. (2009) Institutionalising information asymmetry: governance structures for open innovation, *Information Technology & People*, 22, 4, 297-316.
9. Feller, J., Finnegan, P., Hayes, J., and O'Reilly, P. (2010) Leveraging 'The Crowd': An Exploration of how Solver Brokerages enhance Knowledge Mobility, In *ECIS 2010 Proceedings*.
10. Howe, J. (2006a) The rise of crowdsourcing, Wired magazine, 14, 6.
11. Howe, J. (2006b) Crowdsourcing: A Definition, *Crowdsourcing: Tracking the Rise of the Amateur*.
12. Howe, J. (2009) Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business, Crown Business.
13. Kleemann, F., Voß, G. G., and Rieder, K. (2008) Un (der) paid innovators: The commercial utilization of consumer work through crowdsourcing, *Science, Technology & Innovation Studies*, 4, 1, 5-26.
14. Kozinets, R. V., Hemetsberger, A., and Schau, H. J. (2008) The Wisdom of Consumer Crowds: Collective Innovation in the Age of Networked Marketing, *Journal of Macromarketing*, 28, 4, 339-354.
15. Leimeister, J. M., Huber, M., Bretschneider, U., and Krcmar, H. (2009) Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition, *Journal of Management Information Systems*, 26, 1, 197-224.
16. Malone, T. W., Laubacher, R., and Dellarocas, C. N. (2010) The collective intelligence genome, *MIT Sloan Management Review, Spring*, 51, 3, 21-31.
17. Nickerson, R. C., Muntermann, J., and Varshney, U. (2010) Taxonomy Development in Information Systems: A Literature Survey and Problem Statement, In *Proceedings of the Sixteenth Americas Conference on Information Systems*.

18. Nickerson, R. C., Muntermann, J., Varshney, U., and Isaac, H. (2009) Taxonomy development in information systems: developing a taxonomy of mobile applications, In *Proceedings of the European Conference on Information Systems*.
19. Piller, F. T., Ihl, C., and Vossen, A. (2010) A Typology of Customer Co-Creation in the Innovation Process, *SSRN eLibrary*.
20. Quinn, A. J., and Bederson, B. B. (2009) A taxonomy of distributed human computationCiteseer.
21. Quinn, A. J., and Bederson, B. B. (2011) Human Computation: A Survey and Taxonomy of a Growing Field, In *Proceedings of CHI*.
22. Rouse, A. C. (2010) A preliminary taxonomy of crowdsourcing, In *ACIS 2010 Proceedings*.
23. Ryan, R. M., and Deci, E. L. (2000) Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions., *Contemporary Educational Psychology*, 25, 1, 54-67.
24. Schenk, E., and Guittard, C. (2011) Towards a characterization of crowdsourcing practices, *Journal of Innovation Economics*, 7, 1, 93.
25. Schroer, J., and Hertel, G. (2009) Voluntary engagement in an open web-based encyclopedia: wikipedians and why they do it, *Media Psychology*, 12, 1, 96-120.
26. Sokal, R. R., and Sneath, P. H. A. (1963) Principles of Numerical Taxonomy, Freeman.
27. Whitla, P. (2009) Crowdsourcing and its application in marketing activities, *Contemporary Management Research*, 5, 1, 15-28.
28. Zwass, V. (2010) Co-Creation: Toward a Taxonomy and an Integrated Research Perspective, *International Journal of Electronic Commerce*, 15, 1, 11-48.

Clusters of process types	Crowdsourcing examples with the same process characteristics	Aggregation of contributions	Accessibility of contributions	Remuneration for contributions	Preselection of contributors
Integrative sourcing without remuneration	Delicious, Digg, Facebook Translations, Fashiolista, TripAdvisor, YouTube, Amazon user reviews and ratings	integrative	assess	no	no
	Camclickr, Google Image Labeler, ReCaptcha, Hollywood Stock Exchange	integrative	none	no	no
	Wikipedia, OpenStreetMap	integrative	modify	no	no
	Angie's List	integrative	view	no	no
	eBay reputation system	integrative	view	no	context-specific
	Emporis Community	integrative	modify	no	qualification-based
Selective sourcing without crowd assessment	Netflix Prize, InnoCentive Challenge Center, 99designs (private contests), Brainrack, Calling All Innovators, Crowdspring (private contests), Designenlassen.de (private contests), idea bounty	selective	none	success-based	no
	99designs (public contests), Crowdspring (public contests), Designenlassen.de (public contests)	selective	view	success-based	no
Selective sourcing with crowd assessment	Atizo (Atizo Community), Cisco I-Prize, Threadless	selective	assess	success-based	no
	Atizo (Own Community), InnoCentive@Work	selective	assess	success-based	context-specific
	Dell IdeaStorm	selective	assess	no	no
Integrative sourcing with success-based remuneration	Android Market, Apple AppStore, Yahoo! Contributor Network	integrative	assess	success-based	no
	iStockphoto, YouTube Partners	integrative	assess	success-based	qualification-based
	99designs ready-made logo design	integrative	view	success-based	qualification-based
	Coolspotters	integrative	modify	success-based	no
	Iowa Electronic Markets	integrative	none	success-based	no
Integrative sourcing with fixed remuneration	e-Rewards, Microtask	integrative	none	fixed	context-specific
	LiveOps, Castingwords	integrative	none	fixed	both
	Mechanical Turk	integrative	none	fixed	no

Table 2. Types of crowdsourcing processes