Chapter 2
Background

Crowdsourcing research is a dynamic and vibrant research area, and has been steadily growing over the years.
—Zhao and Zhu (2014)

As crowdsourcing has raised multiple interests, it has been studied in a variety of domains: marketing, management, software engineering, computer science, and information systems. This wide research spectrum enables crowdsourcing to become a young yet rapidly growing field. Publications in this field cover aspects like decision making, quality control, crowd management, workflow design, system architecture and crowd programming (Afuah, Tucci, & Viscusi, 2018; Kohler & Nickel, 2017; Zhao & Zhu, 2014). To help readers understand some key aspects of crowdsourcing, this chapter presents a focused literature review of crowdsourcing research.

The variety of crowdsourcing literature makes the body of knowledge hard to be synthesised. To help achieve a shared structure and understanding of the concept, we propose a layered framework that provides separation of concerns. Figure 2.1 presents the framework comprising of four layers: conceptualisation, classification, process and establishment. These layers are structured symmetrically (top to bottom) from being more abstract to more concrete, and from overview to focus on the research phenomenon.

The first layer conceptualises what crowdsourcing is by characterising three major research streams: crowdsourcing underpinnings, related concepts and existing definitions of crowdsourcing. The literature in each stream is reviewed in Sect. 2.1. The second layer examines the classifications of crowdsourcing and its related elements, which are presented in Sect. 2.2. Classifications and taxonomies are focused because they can provide a structured way to organise knowledge in the field (Nickerson, Varshney, & Muntermann, 2012). Among the different elements classified in the literature, the review highlights the applications, tasks, crowd members and platforms as the most pertinent to this book.

The third and four layers are presented in Sect. 2.3 in order to analyse the current state of business process crowdsourcing. It begins with a review of studies on crowdsourcing processes. The two predominant views, low and high levels of

N. H. Thuan, *Business Process Crowdsourcing*, Progress in IS,
https://doi.org/10.1007/978-3-319-91391-9_2

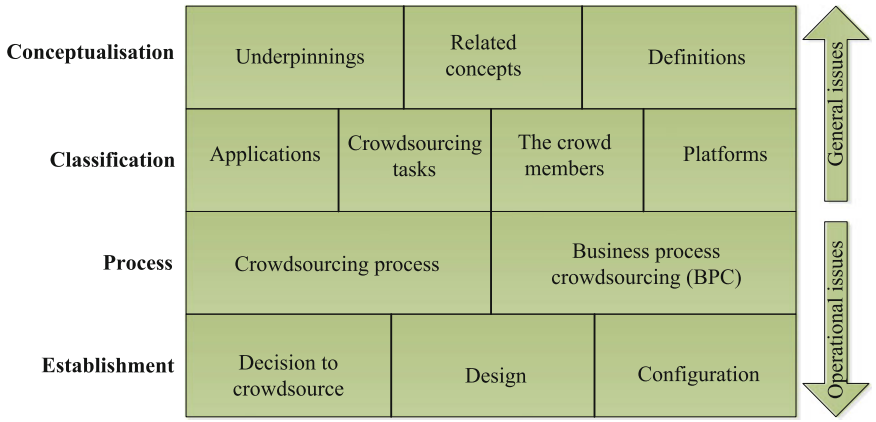

Fig. 2.1 Layered framework for the literature review

granularity for researching crowdsourcing processes, are reviewed. The focus then moves to the more specific concept of BPC. The relatively small body of research related to the concept is reviewed for identifying the important roles of BPC and the need for further investigating BPC. Next, the review analyses the three main stages necessary to establish BPC: decision to crowdsource or not, design process, and configuration. These stages form an analytical theoretical framework providing an abstract picture of BPC and guiding the current research. Altogether, the review provides a comprehensive picture of the current state of business process crowdsourcing.

2.1 The Concept of Crowdsourcing

There is considerable confusion surrounding crowdsourcing terminology in terms of concepts and definitions, as crowdsourcing has continuously developed within different research streams. Thus, it is necessary to explain the concept of crowdsourcing. This section commences with a discussion of the basic ideas behind crowdsourcing. It then compares crowdsourcing with other similar concepts. A definition of crowdsourcing used in this book is then provided.

2.1.1 Main Idea Behind Crowdsourcing

Reliance on the crowd can be traced back to the early 18th century, when the British government decided to provide a cash prize for anyone who could address the problem of precisely calculating ship longitudes (Afuah & Tucci, 2012). Despite a

long history of crowd participation, the concept of crowdsourcing has only really emerged in 2006 when Howe (2006b) introduces a process utilising the crowd for fulfilling Internet tasks. This raises the question why crowdsourcing has become so popular only recently. Investigating this question, three main underpinnings behind the emergence of crowdsourcing have been suggested: (1) the crowd, (2) the organisation, and (3) the medium linking the crowd and the organisation. Let us examine these underpinnings.

First, the crowd's wisdom is one of the main underpinnings enable crowdsourcing (Brabham, 2008a; Hosseini, Shahri, Phalp, Taylor, & Ali, 2015b; Saxton et al., 2013). James Surowiecki calls the underpinning as the 'wisdom of crowds', which claimed that "under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them" (2004, p. xiii). The right circumstances are defined as four prerequisites: cognitive diversity, independence, decentralisation and aggregation. Under these prerequisites, individual ideas in the crowd are not averaged, but aggregated into final solutions. As a result, the aggregated solutions are better than, or at least equal to, the solutions from individual members in the crowd.

Although the wisdom of the crowd is dominant in explaining the concept of crowdsourcing, some extensions should be added to clarify the current capability of crowdsourcing. Malone et al. (2010) extend the underpinnings of crowdsourcing by adding the idea of collective intelligence, which highlights the collective coordination of individuals. This extension opens the solution space of crowdsourcing, based on not only the independence of individuals as the 'wisdom of crowds' but their coordination. Another extension is the ability of crowdsourcing to solve not only single puzzles, but complex tasks that may be decomposed into a large number of simpler tasks (Kittur et al., 2013). As a result, the ability of the crowd should be seen from both its individual and collective intelligence and its capability to manage a large number of tasks.

The second category of underpinnings comes from an organisational viewpoint. As the ability of the crowd seems promising, the next question is whether organisations have any demands for using this ability. In fact, they do. The demands for using external agents to perform tasks has been clearly presented in the management literature: outsourcing (Dibbern, Goles, Hirschheim, & Jayatilaka, 2004), open sourcing (Ågerfalk & Fitzgerald, 2008), and open innovation (Chesbrough, 2013; Seltzer & Mahmoudi, 2013). With outsourcing, organisations have a long history of using contracted resources outside their boundaries. Recently, open sourcing and open innovation have further blurred the organisational boundaries for seeking ideas and innovation beyond the traditional organisational boundaries.

The demands for external sourcing explains the reason why organisations have largely been attracted to crowdsourcing. Organisations utilising the crowd may get benefits similar to outsourcing and open innovation, such as cost saving, customer involvement, and access to outside skills (Rouse, 2010; Saxton et al., 2013). Further, crowdsourcing allows organisations to leverage flexible, on-demand labour. These benefits increase organisational demands for crowdsourcing. It is important to note that although organisational demands to use external resources of

crowdsourcing are similar to outsourcing, open sourcing, and open innovation, these concepts are distinctive because of other characteristics, as discussed in the next section.

Given the aforementioned underpinnings, the term 'crowdsourcing' can be etymologically analysed as a combination of two words: *crowd* and *sourcing*. However, the fact that these underpinnings have existed long before the recent emergence of crowdsourcing reveals that another underpinning is needed to enable crowdsourcing. Most of the crowdsourcing literature agrees on the role of the Internet, and in particular the recent dominance of Web 2.0 (Brabham, 2013; Saxton et al., 2013; Zhao & Zhu, 2014). Being globally collaborative, Web 2.0 has changed the nature of online interaction where individuals are no longer passive receivers but active contributors (Brabham, 2013; OReilly, 2007). Brabham (2013) notes that Web 2.0 fastens a voluntary participatory culture onto a global, virtual environment, where Internet users are willing to contribute their skills and labour. Such contributions are perceived as valuable resources for work.

Further, Web 2.0 empowers the open call, which is a distinctive characteristic of crowdsourcing. Because of its millions of users, Web 2.0 extends the scope of the open calls through providing a valuable medium for approaching innumerable anonymous audiences (Saxton et al., 2013). In other words, any given interested participants can now participant in crowdsourcing. It has also eased users to participate in a variety of Internet activities with fewer barriers, e.g. regarding time and space (Brabham, 2013). As a result, it extends the reach and the scope of the crowdsourcing open calls.

This review has shown that, the combination of the crowd, Web 2.0, and organisational demands, can explain the emergence and foundations of crowdsourcing. Given these underpinnings, the IS discipline, which is concerned with people, technologies, and organisations (Bacon & Fitzgerald, 2001), has crowdsourcing as a focus point. This focus point also comes from a strength of the IS research, which draws upon reference disciplines to build its own knowledge base (Baskerville & Myers, 2002). This is exactly the need for the field of crowdsourcing, as a large part of research into crowdsourcing is not very well delimited. All in all, we believe that IS research like the current research can make significant contributions to progress the crowdsourcing field.

The review has also shown that no single underpinning can enable crowdsourcing per se, but rather the combination of the three underpinnings supports the emergence of crowdsourcing. This combination distinguishes crowdsourcing from other concepts, being presented in the next section.

2.1.2 Related Concepts

In another stream of research attempting to clarify the concept of crowdsourcing, many researchers compare this notion with closely related concepts, such as open innovation, outsourcing, open source, and peer production. This section reviews

this research stream and discusses crowdsourcing by comparing its similarities and differences with the related concepts.

Among the competing concepts, one often discussed in relation to crowdsourcing is *open innovation*. Crowdsourcing and open innovation share a common basis where organisations embrace openness to harvest external knowledge and expertise, the opposite of closed innovation. As a result, some researchers suggest that crowdsourcing belongs to or is a technique of open innovation (Marjanovic, Fry, & Chataway, 2012; Seltzer & Mahmoudi, 2013). However, other researchers argue that these two concepts are different, at least in two important points. First, open innovation mainly focuses on innovation processes, while crowdsourcing has been used for various types of tasks (Nakatsu, Grossman, & Iacovou, 2014; Schenk & Guittard, 2011). Second, organisations interact mainly with other firms and their stakeholders in open innovation, but rely on anonymous crowd members in crowdsourcing activities (Flostrand, 2017; Schenk & Guittard, 2009).

Outsourcing is another concept closely related to crowdsourcing. As noted in the previous section, the two concepts are similar on the organisational demands for external agents. As a result, pioneering researchers considered crowdsourcing as a form of outsourcing (Howe, 2006b; Rouse, 2010; Whitla, 2009). Nevertheless, recent conceptualisations of crowdsourcing clearly identify the differences between these two concepts. One major difference is who performs the activities. Actors performing crowdsourcing tasks are informal members of the crowd, while in outsourcing they are mainly established supplier firms. Another difference lies in how to manage these actors. Compared to the official contracts used in outsourcing, crowdsourcing uses an open call where any member in the crowd can participate in the project (Zhao & Zhu, 2014). Finally, financial incentives are the main motivation for task performers in outsourcing, whereas crowdsourcing can be based on both intrinsic incentive, e.g. personal enjoyment and hobby, and extrinsic incentives, e.g. money (Hossain, 2012; Kaufmann, Schulze, & Veit, 2011; Naderi, 2018).

The literature also distinguishes crowdsourcing from *open source*, although the two concepts are based on resources from the community to accomplish tasks. There are two key aspects distinguishing them: management and engagement. In crowdsourcing, activities are managed by the organisations, whereas in open source these activities are self-managed and community-driven (Brabham, 2013). Regarding to how the community is engaged to perform the activities, crowdsourcing outcomes can be achieved either independently or collaboratively (Geiger, Seedorf, Schulze, Nickerson, & Schader, 2011), but outcomes from open source are achieved mainly through collaboration. The motivation of the community is another difference between these two concepts. Most of the time, members in open source communities perform tasks based on intrinsic motivation, whereas both intrinsic and extrinsic motivations can be found in crowdsourcing (Kaufmann et al., 2011; Naderi, 2018). Furthermore, unlike open source, crowdsourcing campaigns clearly have intellectual property rights and are not restricted to software development (Wu, Tsai, & Li, 2013).

A few researchers equate crowdsourcing to a form of *peer production* (Mason & Watts, 2009; Wu, Wilkinson, & Huberman, 2009). These researchers believe that

peer production sites, like YouTube, can be seen as crowdsourcing because contents on these sites are created by mass individuals in the crowd. However, other researchers argue that crowdsourcing is completely different from peer production. Estellés-Arolas and González-Ladrón-de-Guevara (2012) suggest that crowdsourcing tasks require clear objectives, and thus YouTube, where an individual can upload any video, is not crowdsourcing. In addition, peer production mainly depends on particular communities (Haythornthwaite, 2009; Huberman, Romero, & Wu, 2009), whereas crowdsourcing relies on anonymous members of the crowd, as previously mentioned.

To summarise the above discussion, this review adapts Malone et al.'s (2010) framework to compare crowdsourcing with the related concepts. This framework includes four questions: what needs to be performed, who is performing the task, why people do this, and how the task is being done. An additional question about controlling intellectual property (IP) is added for clarifying the locus of control on the outcomes. By answering the five questions (five rows), Table 2.1 presents the main differences between crowdsourcing and the other concepts. This table reflects that crowdsourcing is a distinctive notion, leading us to investigate the concept per se.

2.1.3 Crowdsourcing Definition

Given the different concepts related to crowdsourcing, we are not surprising that researchers have defined the crowdsourcing concept differently. This section presents a brief history of crowdsourcing definitions in order to understand the concept evolution, and ultimately to form a definition for use in this book.

Until now, crowdsourcing has a short history of one decade. The phenomenon began to appear in 2006 after Howe (2006b) coined this term when he observed several websites utilising Internet users to perform certain activities. It is interesting to note that Howe's (2006b) article has appeared in Wired Magazine—a news media, which indicates that crowdsourcing is a concept spreading from practice to academia. In the article, crowdsourcing was described as the act of organisations through the form of an open call in order to "tap the latent talent of the crowd" (Howe, 2006b, p. 2). In the same year, he proposed the first definition of crowdsourcing.

> Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers (Howe, 2006a).

Up to now, this definition is among the ones most cited in the field due to its exploratory nature and simplicity. It is worth noting two interesting points from this definition. First, it views organisations as the main caller who operationalise

Table 2.1 Main differences between crowdsourcing and related concepts

	Open innovation	Outsourcing	Open source	Peer production	Crowdsourcing
Tasks	• Only innovation		• Software	• Undefined tasks	• Varied types of tasks • Predefined tasks
Workforce	• Other firms and customers	• Supplier firms	• Software community	• Certain community	• Members of the crowd
Participant motivation		• Extrinsic motivations	• Intrinsic motivations		• Intrinsic and extrinsic motivations
Nature of management and engagement		• Official contracts	• Workflows and quality control managed by community • Collaborative	Collaborative	• Open call • Without official contract • Workflows and quality control mainly managed by the organisations • Collaborative and independent
Control on IP			• IP open		• IP protected

crowdsourcing, which is completely aligned with the promotion of crowdsourcing for organisations in the book. Second, in this definition, crowdsourcing is a sourcing strategy and is an extension of outsourcing.

After 2006, researchers started to explore crowdsourcing and soon published several alternative definitions. Since then, crowdsourcing definitions have evolved over time. Figure 2.2 summarises the evolution of crowdsourcing definitions during the last decade.

After Howe's (2006a) definition, several academic definitions of the concept were published between 2008 and 2009. Extending Howe's (2006a) view, some researchers conceptualised crowdsourcing as a sourcing model where the task performers were the crowd. These researchers further defined who the crowd was and positioned it as a workforce alternative to internal employees and outsourcing agents (Ågerfalk & Fitzgerald, 2008; Howe, 2008; Whitla, 2009; Yang, Adamic, & Ackerman, 2008). At the same time, a parallel approach focused on the intelligence capabilities of crowdsourcing. Researchers in this approach defined crowdsourcing as a problem solving model, where the crowd contributes not only with labour but also with creativity (Brabham, 2008a, 2008b; DiPalantino & Vojnovic, 2009; Vukovic, 2009). As a pioneer researcher in this stream, Brabham (2008b) summarised the notion of crowdsourcing as "a process, a model, for distributed problem solving through the Web" (p. 1). The term 'problem' in Brabham's definition

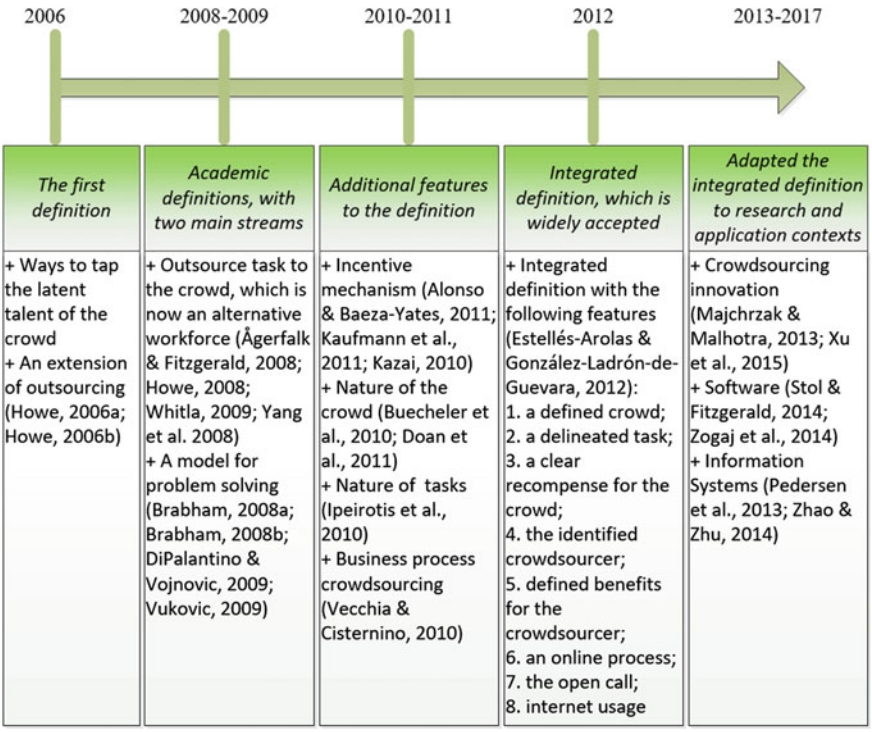

Fig. 2.2 Evolution of crowdsourcing definitions

should be understood in a broad sense, including not only R&D problems but also design and innovation.

During 2010–2011, there was a boom of crowdsourcing definitions, aligning to a diverse set of practices and an increasing number of research interests in the field. At this stage, researchers adopted different theoretical bases and models to investigate several aspects of crowdsourcing. Depending on the research foci, the related features were depicted and added to crowdsourcing definitions, including the nature of the crowd (Buecheler, Sieg, Füchslin, & Pfeifer, 2010; Doan, Ramakrishnan, & Halevy, 2011), the nature of tasks (Ipeirotis, Provost, & Wang, 2010), and incentive mechanisms (Alonso & Baeza-Yates, 2011; Kaufmann et al., 2011; Kazai, 2010). These definitions, on the one hand, contribute to clarifying several features of the concept. On the other hand, definitions with too many additional features suffer from diversity and sometimes conflict with each other, which makes crowdsourcing hard to comprehend.

Addressing this problem, in 2012, Estellés-Arolas and González-Ladrón-de-Guevara (2012) aimed at establishing an integrated definition of crowdsourcing. Driving their research was the goal to conceptualise 'any given crowdsourcing activity' by reviewing the diverse definitions extracted from literature. The authors selected 209 crowdsourcing articles and analysed 40 of them that present original definitions of crowdsourcing. The

results suggest eight key characteristics of crowdsourcing: a clearly defined crowd, a task with a clear goal, a clear recompense for the crowd, an identified crowdsourcer (or caller), defined compensation for the crowdsourcer, an online process, an open call, and Internet usage. The authors then integrate these characteristics into a single comprehensive definition.

> Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organisation, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken (Estellés-Arolas & González-Ladrón-de-Guevara, 2012, p. 197).

Due to its comprehensiveness, this definition has been widely accepted and frequently referred to. Yet, the definition is wordy and thus complex, which decreases its practical use. As a result, many recent studies have implicitly or explicitly adapted the aforementioned eight characteristics by simplifying and adjusting them to their own research and application contexts. For instance, crowdsourcing software emphasises the software tasks being crowdsourced (Stol & Fitzgerald, 2014) and the intermediated platforms (Zogaj, Bretschneider, & Leimeister, 2014); crowdsourcing innovation focuses on the innovative ability of the crowd (Majchrzak & Malhotra, 2013; Xu, Ribeiro-Soriano, & Gonzalez-Garcia, 2015). These adaptations show that there is no unique universal definition of crowdsourcing appropriate for all applications and research contexts, though Estellés-Arolas and González-Ladrón-de-Guevara's definition may form a basic understanding.

Aligning to the most recent trend, this book simplifies and adapts the definition by Estellés-Arolas and González-Ladrón-de-Guevara (2012) to the organisational context. We use the following definition.

> Crowdsourcing is an online strategy in which an organisation proposes defined task(s) to the members of the crowd via a flexible open call. By undertaking the task(s), the members contribute their work, knowledge, skills and/or experience and receive rewards, including economic rewards, social recognition, self-esteem, or the development of individual skills. The organisation will obtain contributions from the crowd and will utilise the results to meet business goals.

To sum up, this review has shown the conceptualisation of crowdsourcing, through three facets. The first facet has shown the three main pillars of crowdsourcing: the organisational demands for external sourcing; the ability of the crowd; and the intermediary Web 2.0. These pillars together enable crowdsourcing. The second facet has compared and differentiated crowdsourcing with related concepts, like open innovation, outsourcing, open sourcing and peer production. It emphasises the distinctive characteristics of crowdsourcing. The final facet has shown a brief evolution in crowdsourcing definitions. It then proposes the definition that to

be used in this book. From these facets, it is worth noting that although a few referencing theories have been applied to explain crowdsourcing, like the wisdom of the crowd, open innovation, and outsourcing practices, the distinctive characteristics of crowdsourcing state that these theories are not predominant in the phenomenon. Consequently, crowdsourcing is a concept per se that needs its own structures and establishment.

2.2 Crowdsourcing Classifications, Taxonomies, and Typologies

As classifications and taxonomies are useful to structure knowledge in the IS discipline (Nickerson et al., 2012), an extensive body of crowdsourcing literature is devoted to crowdsourcing classifications, taxonomies, and typologies. Although they contribute to structuring the domain, each of them focuses on different crowdsourcing elements. This section considers the popular classified elements: applications, tasks, members and platforms, which are essential for exploring crowdsourcing processes. In particular, this section aims to answer the following four questions: What are the crowdsourcing applications? Which types of tasks can be crowdsourced? Who will perform these tasks? And where can these tasks be performed?

Before proceeding, it is useful to clarify terminologies of classification, as a variety of them have been used in the literature, including classification, taxonomy, and typology. The term *classification* has been used as both a product and a process of classifying objects according to a particular system (Fettke & Loos, 2003). *Taxonomy* and *typology* are two forms of classification that usually deploy multi-dimensions to classify objects into categories. Some researchers further distinguish taxonomies as empirical classifications and typologies as conceptual classifications (Bailey, 1994). However, other researchers suggest using classification, taxonomy, and typology interchangeably (Gregor, 2006; Nickerson et al., 2012). We follow this suggestion as we observe that the crowdsourcing literature commonly refers to the three terms in an interchangeable way. Consequently, this book uses these terminologies more or less synonymously.

2.2.1 Applications

Crowdsourcing has been applied to different *applications*. Howe (2006b) discusses the crowdsourcing concept through several applications in solving real business problems, including InnoCentive for problem solving, iStockphoto for image exchange, and AMT for micro tasks. In addition to business applications, crowdsourcing can also be applied to scientific research, urban planning, public health,

and cultural heritage. Given the variety of crowdsourcing applications, their classifications are necessary for understanding the potential utility of crowdsourcing. Addressing this necessity, several application classifications have been proposed in the literature, which are now reviewed in detail. The review moves from simple classifications, defining for specific areas, to more inclusive typologies at the level of work practice.

Some studies, maybe for simplification, classify crowdsourcing applications specifically for one single area. For instance, Whitla (2009), focusing on marketing-related areas, classifies crowdsourcing applications into three function categories, namely marketing research, product development, and advertising and promotion. Gomes et al. (2012) propose a crowdsourcing taxonomy with a focus on the context of musical productions. Based on what crowdsourcing can be utilised for, the taxonomy identifies six types of applications: music co-creation, decision support, crowdsourced music collection and management, promoting music information, market place, and crowd funding.

Adopting a broader approach, other studies propose a number of application typologies that can be used in multiple domains. Kleeman et al. (2008) explored start-up crowdsourcing applications, and typologically grouped them according to their functions. As a result, seven application types are defined, namely product development and configuration, product design, permanent open calls, competitive bids, community reporting, product rating, and customer-to-customer support. This approach is also employed by Brabham in his recent book, *Crowdsourcing* (2013). He, surveying crowdsourcing cases, conceptualises them into four different functions, including knowledge discovery and management, broadcast search, peer-vetted creative production, and human intelligence tasks. Other typologies which follow a similar approach are mentioned in the literature (Man-Ching, King, & Kwong-Sak, 2011).

The studies reviewed so far have a common point. They suggest that *function* is a main dimension to classify crowdsourcing applications. Agreeing with this suggestion, we note, however, that functions alone seem not enough, since a *context*, where crowdsourcing is applied, plays an equally significant role. Chandler and Kapelner (2013), who conducted an experiment on AMT, find that if the context is explained, more workers are willing to participate in the crowdsourcing application. In addition, whether it is a business or non-business context strongly influences the application operation because the context directly links to incentives that may be required to attract people to participate in the crowdsourcing applications (Rosen, 2011).

Given the importance of contexts in characterising crowdsourcing applications, two dimensions: *function and context* together are likely more appropriate to classify applications. This appropriateness is supported by Zhao and Zhu (2014), who broadly reviewed crowdsourcing applications. By deductively analysing 126 applications, they propose a typology based on the two dimensions of function and context. In the first dimension, these authors group functions into four categories: design and development, test and evaluation, idea and consultant, and others. In the second dimension, two categories of contexts are suggested: business and non-business. A business context consists of for-profit organisations, while

Table 2.2 Typology of crowdsourcing applications (Zhao & Zhu, 2014)

Context	Function			
	Design and development	Idea and consultant	Test and evaluation	Other
Business	• Threadless • IStockphoto	• MyStarbucks Idea • InnoCentive	• Crowdspirit	• AMT
Non-business	• NextStopDesign (Brabham, 2012)	• QuestVille	• UTest	• Wikipedia

non-business includes non-profit organisations and institutions (Zhao & Zhu, 2014). Although this dimension considers contexts at an organisational level, we suggest the level of application is more precise for this dimension. The reason is that one organisation may have both business and non-business applications, such as Amazon owning AMT for profit and QuestVille for non-profit (Saxton et al., 2013). In this case, the context dimension does not associate with the organisation but with its applications. Therefore, this book adopts the typology proposed by Zhao and Zhu (2014), yet considers both the function and context dimensions from the viewpoint of crowdsourcing applications (Table 2.2).

2.2.2 Tasks

Tasks are basic elements of a crowdsourcing application. Organisations define tasks and send them to members in the crowd, who will perform these tasks. Several studies have suggested clearly identifying task characteristics before crowdsourcing, which helps to determine the appropriate approach for a particular task (Malone et al., 2010; Nakatsu et al., 2014; Rosen, 2011). Several taxonomies characterising tasks have been proposed in the literature.

There are two main views on building task taxonomies regarding whether tasks should be examined in related with other elements or by its own nature. On the one hand, a number of published taxonomies are based not only on task properties, but also on "key questions [elements] associated with a single task" (Malone et al., 2010, p. 22). Rouse (2010) provides one the of the earliest taxonomies, structured around three dimensions: nature of the task, distribution of benefits, and forms of motivation. In a similar vein, Malone et al. (2010) propose a multi-dimensional classification after analysing 250 instances of crowdsourcing. The classification is based on four basic questions: what is being crowdsourced, who is performing the task, why would people do this, and how is the task to be done. In these cases, the developed taxonomies suggest multiple dimensions for classification, with task as a central dimension.

On the other hand, another group of published taxonomies classifies tasks by their own nature. By examining the task characteristics in practical applications,

Table 2.3 Examples of crowdsourcing task types (adapted from Schenk and Guittard (2011))

Complexity	Participation mode	
	Individual (integrative)	Competitive (selective)
Simple	**Market place** • Simple tasks (MicroWorkers, AMT & Taskcn)	**Simple contest** • Answering simple questions (Ask Ville by Amazon & Yahoo Answers)
Skilled	**Collective intelligence** • Writing & editing (Wikipedia) • Writing academic papers (Tomlinson et al., 2012)	**Problem solving contest** • Designing T-shirts (Threadless) • Problem solving (InnoCentive)

Schenk and Guittard (2009, 2011) suggest two dimensions to classify crowdsourcing tasks. The first dimension classifies tasks as simple, complex or creative. Simple tasks are jobs that can be performed without any specific skills, such as text transcription. Complex tasks require expertise and skills, such as problem solving. Creative tasks relate to individual creativity, such as logo design. The second dimension distinguishes between the integrative and selective nature of tasks (Schenk & Guittard, 2011). Other taxonomies in this group can also be found in work by Nakatsu et al. (2014).

Given the existing taxonomies, a critical question is which one will be used in this book. To answer this question, the book adopts Nickerson et al.'s (2012) suggestion that usefulness is the key criterion to evaluate a taxonomy and its dimensions. Thus, choosing dimensions for task classification in the book should be based on their usefulness for the research focus. That is, the establishment of BPC, consisting of three stages: the decision to crowdsource, process design, and configuration, will be discussed in Sect. 2.3.3. In the first stage, the complexity of tasks plays a role in the decision to crowdsource (Zhao & Zhu, 2014). In the remaining stages, whether tasks are achieved individually or competitively, influences the crowdsourcing design and operation, because it directly affects how the tasks should be planned, coordinated, and performed.

Consequently, this study adapts the two dimensions proposed by Schenk and Guittard (2011): *task complexity* (simple and skilled) and *the difference between integration and selection based crowdsourcing*. Table 2.3 presents examples of different types of crowdsourcing tasks (and their related platforms).

2.2.3 Members of the Crowd

Crowd members are actors who accomplish tasks in crowdsourcing applications. There are several studies examining characteristics of crowd members. In general, these studies can be grouped into one of two research directions. The first direction examines the crowd characteristics by exploring its properties, such as who members of the crowd are and where they come from. Studies by Mason and Suri

(2012) and Brabham (2011) can be categorised in this direction. Another direction studies the crowd as a whole and assesses its performance (Chandler & Kapelner, 2013; Stewart, Lubensky, & Huerta, 2010).

In the first direction, Brabham (2011) changed the popular image of the crowd being amateur. By conducting a survey on iStockphoto and several interviews on Threadless, he finds that members on both of these platforms "seem ill-fitted to the amateur label" (Brabham, 2011, p. 399). Specifically, 47% of participants on IStockphoto described themselves as professional, while many members on Threadless have previously performed real design activities (Brabham, 2011). The argument that the crowd is not wholly amateur, and thus can be in competition with professionals, is also supported by other studies. Jeppesen and Lakhani (2010), who examined the members on InnoCentive, report that "65% of solvers reported holding Ph.D. degrees" (p. 1026). Poetz and Schreier (2012), conducting a case study in the idea contest, find that the crowd can outperform the professionals in certain aspects of idea quality.

Similar to Brabham (2011) in exploring the properties of the crowd, Mason and Suri (2012) focusing on AMT present several aspects of AMT's 'workers'. For instance, there are about 100,000 workers on AMT, who are mainly from USA and India. This crowd has more females than males. These characteristics are consistent with findings from another study of 1,000 workers using the same platform conducted by Paolacci et al. (2010), who further report that the hourly average wage on this platform is $1.66. From these observations, three reasons provided by Mason and Suri (2012) to choose AMT for online experimentation can be generalised as the crowd characteristics on AMT: large pool of workers, pool diversity and low cost.

In the other direction, studies investigating performance of the crowd as a whole show that the performance is not as promising as the characteristics presented in the first direction. The fact that not all members of the crowd actively performed tasks was analysed by Stewart et al. (2010), building on the participation inequality rule of online community (Nielsen, 2006). By analysing a crowd of 400,000 members in a language translation application, these authors separate members of the crowd into three categories: super contributors (1%) who provide the most contributions, contributors (66%) who provide the moderate contributions and outliers who rarely contribute (33%). Further analysing the crowd members, Kazai et al. (2011) find that members may perform tasks dishonestly, randomly, or sloppily. In a similar vein, Vuurens and de Vries (2012) suggest a theoretical typology classifying four types of workers regarding their behaviours: diligent workers, sloppy workers, random spammers, and uniform spammers.

From the given discussion, some characteristics of the crowd should be highlighted. On the one hand, the crowd is promising in terms of providing a large, diverse, and low-cost workforce (Mason & Suri, 2012). It may also include 'self-selected' experts (Brabham, 2011). On the other hand, members of the crowd have different levels of contribution for accomplishing tasks (Stewart et al., 2010; Vuurens & De Vries, 2012). We note that the reviewed studies mainly identify the crowd characteristics based on individual applications and platforms, which implies

that the characteristics of the crowd may be different in varied applications and contexts.

2.2.4 Platforms

Platform is another key element of crowdsourcing, which serves as a mediator connecting the organisation and the crowd (Hirth, Hoßfeld, & Tran-Gia, 2011). Vukovic (2009) describes several functions of a crowdsourcing platform: "issues authentication credentials for requestors and providers when they join the platform, stores details about skill-set, history of completed requests, handles charging and payments, and manages platform misuse" (p. 687). Aiming to utilise crowdsourcing, organisations can choose either to develop their own platforms or to use the available ones provided by a third party. Each approach has its own advantages and disadvantages.

Some examples of organisations developing their own crowdsourcing platforms are Threadless, and MyStarbucksIdea in the business context, and Next Stop Design in the non-business context (Brabham, 2012). Through self-development, organisations can fully control the application and its functions, such as tracking geographic locations of visitors for research purposes in case of Next Stop Design (Brabham, 2012). Another advantage of this approach is building closer relationships with their own customers, who associate with the platforms. For instance, Threadless uses a self-developed platform to ask customers to design T-shirts, and then sells those T-shirts to the customers (Brabham, 2010). Despite these advantages, this approach requires organisations having experts and experience in developing crowdsourcing platforms, since the platform development may have several complex requirements (Adepetu, Ahmed, Al Abd, Al Zaabi, & Svetinovic, 2012; Vukovic, 2009).

As an alternative to self-development, organisations can hire existing crowdsourcing platforms built by a third party to deploy their applications. The existing platforms can be further divided into two kinds: specialised platforms, which focus on particular tasks (Hirth et al., 2011; Hoßfeld et al., 2013); and horizontal platforms, which can be utilised for different types of tasks (Kucherbaev et al., 2013). Two examples of a specialised platform are InnoCentive that utilise the crowd only for problem solving purposes (Malone et al., 2010), and TopCoder that uses crowdsourcing for software engineering (Mao, Capra, Harman, & Jia, 2017). Differently, horizontal platforms publish different types of tasks. AMT is a typical horizontal platform, which can help an organisation to do several tasks, including data collection, transcription, and image categorisation. To configure a crowdsourcing application on horizontal platforms like AMT, organisations need to use the provided application programming interface (API) (Ipeirotis et al., 2010). Thus, basic programming skills and platform knowledge are required.

Using existing platforms can save organisations' resources, which would otherwise need to be spent on developing their own new platform. Furthermore,

existing platforms already have their own members, who are available for new crowdsourcing applications. For instance, an application developed on AMT can use any number of 100,000 available workers (Mason & Suri, 2012). However, existing platforms limit crowdsourcing applications to what is supported by the platforms. From the above discussion, it is important to note that both approaches have their own pros and cons, which should be considered when making the decision to build or to hire a crowdsourcing platform. Table 2.4 summarises the main pros and cons of the discussed platforms types: self-development, specialised platforms, and horizontal platforms.

In summary, the preceding review identified major classifications in the crowdsourcing literature, including applications, tasks, members, and platforms. On the one hand, these classifications suggest possible options and features that are available in crowdsourcing, which contributes to initially structure the domain. On the other hand, many of them have focused on specific aspects of crowdsourcing and on specific crowdsourcing contexts. This leads to differences, sometimes conflicting, on the domain structures. For instance, crowdsourcing tasks can be classified differently using either four dimensions (Malone et al., 2010), three dimensions (Rouse, 2010), or two dimensions (Schenk & Guittard, 2011).

We believe that this is symptomatic of a more general issue with the ad hoc focus of the existing classifications. That is, the domain is structured through its individual elements without synthesis and coordination between them. If we cannot address this ad hoc issue, and if new studies continue to propose crowdsourcing taxonomies that are solely relevant to specific elements, the domain may end up with ambiguity over its structure. Given that, there is a strong need for a more comprehensive integrated approach in order to structure the domain.

Addressing the need, we suggest that a domain ontology and a process view are necessary for structuring the domain. Regarding the former, a domain ontology

Table 2.4 Crowdsourcing platform types

Dimension	**Self-development platforms**	**Platforms by a third party**	
		Specialised platforms	Horizontal platforms
Control	Fully control	Depending on platform	
Customer relationship	High	Low	
Development effort	High	Low	
Tasks being crowdsourced	Organisational focus	Platform focus	Diversity
Availability of crowd	Low	Medium	High
Crowd expertise	High	High	Low
Examples of platforms	• MyStarbucksIdea • Next stop design (Brabham, 2012)	• InnoCentive • TopCoder	• AMT • Microworkers

enables us to integrate the existing classifications. Nickerson et al. (2012) suggest that ontologies are the next stage of taxonomy development. Further, Corcho et al. (2003) highlight ontologies for their comprehensiveness and ability to structure domain knowledge. Regarding the need for a process view, we note that the existing classifications have not been linked together yet, which is necessary to constitute the whole crowdsourcing practice. This highlights the process view connecting individual elements in a meaningful way. This process view is a central of the book, where we address crowdsourcing processes and BPC, and is the focus of the next section.

2.3 Current State of Business Process Crowdsourcing

This section aims to paint an overall picture regarding the emerging state of business process crowdsourcing (BPC). The section starts with describing crowdsourcing processes, an antecedent of BPC. It then provides a review of BPC related literature, followed by an initial conceptualisation and a theoretical framework of BPC. By channelling the related research, the framework guides the current research and paints an abstract picture of BPC.

2.3.1 Crowdsourcing Process

The notion of a *crowdsourcing process* is critical to operationalise a crowdsourcing strategy. Thus, it is a recurrent topic in the crowdsourcing literature. We use the term 'process' to refers to a set of systematic activities to complete some deliberate results. Well-coordinated processes are assumed not only to generate better crowdsourcing results (Thuan et al., 2017), but also to deploy crowdsourcing applications more efficient and with less cost (Tranquillini et al., 2015). Numerous studies have devoted attention to the topic. By and large, existing studies on crowdsourcing processes can be classified into two basic genres according to its view: high and low levels of granularity.

With high level of granularity, some studies adopt a holistic view to conceptualise the crowdsourcing process. Early, research referred to crowdsourcing processes with an understanding purpose. Consequently, crowdsourcing processes were conceptualised by rich descriptions with several illustrative examples (Leimeister, Huber, Bretschneider, & Krcmar, 2009; Whitla, 2009), and by identification and description of actions executed by different crowdsourcing actors (Geiger et al., 2011; Vukovic, 2009; Wexler, 2011). At this early time, crowdsourcing processes were mostly studied together with other foci like crowdsourcing applications and taxonomies, rather than as a separate primary research focus. Before moving to review studies that primarily investigate crowdsourcing processes, we synthesise the existing descriptions to provide a narrative sketch of the

crowdsourcing process. More precisely, we adapt the earliest but most widely used description by Whitla (2009) and add into it supplementary descriptions. As a result, a crowdsourcing process can be described as follows.

The crowdsourcing process starts with a go/no-go decision whether to choose crowdsourcing to perform the organisational tasks or not (Thuan, Antunes, & Johnstone, 2013; Wexler, 2011). If the decision to crowdsource is made, the organisation then creates an open call to release the defined tasks to the crowd. This step is normally done through a platform developed by either a third party or the organisation itself. Through the open call, the organisation approaches members of the crowd, who can belong specifically to a particular community or just anyone willing to complete the task. The members accomplish these tasks individually or collaboratively, and then submit the results back to the organisation, which assesses the quality of the results. Incentives will be given to the members if the organisation is satisfied with the submission results (Whitla, 2009). The results are intended to be incorporated into organisational activities (Leimeister et al., 2009; Wexler, 2011).

Keeping in mind the initial descriptions, researchers started to explore crowdsourcing processes from a high level of granularity. Aiming to identify the main structures of the process, they commonly adopted an abstract view to discover the main stages and concerns in the process. Brabham (2009, 2012), exploring a crowdsourcing project for public participation in transit planning, formulates a crowdsourcing process using four stages. First, a problem that needs to be solved and its related information are clarified. Second, an open call is sent to the crowd through a self-developed website. This call includes data necessary to solve the problem, reward information and the intended format of the solutions. Third, crowd members can choose to participate in the project. Finally, the organisation evaluates the proposed solutions to choose the winners.

Also adopting an abstract broad view, Muhdi et al. (2011) conducted an explorative case study to analyse twelve crowdsourcing projects. As a result, they formulate the main operations in the crowdsourcing process as five stages: deliberation, preparation, execution, assessment, and post-processing. In the first stage, organisations analyse crowdsourcing and "decide whether the crowdsourcing approach is appropriate to solve their internal problem[s]" (p. 322). If the decision to crowdsource is made, the second stage involves choosing a particular platform that is appropriate for the crowdsourcing activity. The next two stages are dedicated to executing the crowdsourcing activity on the chosen platform, and evaluating the received results. The final stage transfers the received results, such as ideas and solutions, to real organisational implementation.

In a similar vein, Stol and Fitzgerald (2014) conducting case study research recently examined crowdsourcing processes in the context of software companies. However, they structure their findings differently compared to the two aforementioned studies. More precisely, instead of formulating crowdsourcing processes as a set of sequential stages, they identify major building blocks of the crowdsourcing process, including task decomposition, coordination and communication, planning and scheduling, quality assurance, managing knowledge and intellectual property, and providing incentives to the crowd. Similar approaches that formulate main

elements of the crowdsourcing process by case study are quite common (Ågerfalk, Fitzgerald, & Stol, 2015; Zogaj et al., 2014).

Overall, this group of studies views the crowdsourcing process as an important research focus and contributes empirical efforts to formulate the main stages and building blocks that comprise the crowdsourcing process. However, the main research methods adopted in these studies are exploratory case studies (Ågerfalk et al., 2015; Muhdi et al., 2011; Stol & Fitzgerald, 2014; Sutherlin, 2013). The exploratory nature, together with the particular investigated cases/contexts, explains the existence of different, likely one-off crowdsourcing processes. Furthermore, as this group of studies target to provide an overall picture of the crowdsourcing process, they focus on high-level abstract concepts and thus face significant gaps mapping the abstract concepts to specific workflows or activities, necessary to establish the crowdsourcing process.

With low level of granularity, a large number of studies have investigated varied aspects of the crowdsourcing process. Although they have helped specify workflows and activities necessary to establish the crowdsourcing process, their ad hoc nature has been repeatedly complained (Geiger & Schader, 2014; Zhao & Zhu, 2014). This ad hoc nature is further revealed through two aspects. First, different research methods have been adopted to examine the specific activities of the crowdsourcing process. For instance, methods for researching task definition include lab experiments (Khazankin et al., 2012a), open-ended and quantitative surveys (Schulze, Seedorf, Geiger, Kaufmann, & Schader, 2011), and engineering design (Bozzon, Brambilla, Ceri, & Mauri, 2013). These differences contribute to clarifying different aspects of the activity, yet a comprehensive approach is still missing. Second, the domain is lacking a strong knowledge base guiding crowdsourcing process establishment (Palacios, Martinez-Corral, Nisar, & Grijalvo, 2016; Zhao & Zhu, 2014). As a result, the domain knowledge remains scattered, varied and sometimes conflicting.

Given the existence of the large number of studies in this group, this section does not intend to review them one by one, which will be the focus of the scoping knowledge source in Sect. 3.1. Rather, we summarise other major literature reviews, which characterise the complexity and isolated concerns of the crowdsourcing field. Among a few literature reviews in the domain, we focus on the two most recent and major reviews.

In 2014, the first major review was published by Zhao and Zhu (2014). These authors identified 55 crowdsourcing papers, based on a systematic search and selection of all major scholar databases in the period from 2006 to 2011. Analysing the papers, they suggest that "empirical studies have been conducted almost entirely on events/processes" (p. 419). These authors further map these ad hoc foci into major themes, and outline future research directions, including motivation to participate, participant's behaviour, making the decision to adopt crowdsourcing, governance and implementation, quality control and evaluation, incentive mechanisms, and technological issues, which are all major topics of crowdsourcing process studies. The review also indicates the emerging nature of the domain because only a small part of the studies (16%) is based on theoretical foundations.

Amrollahi (2015), among the most recent reviewers, aims at synthesising the crowdsourcing literature into a process model. He started by searching crowdsourcing papers in the period of 2009 to early 2014, coming up with 566 papers, and then selected 39 papers directly focusing on the crowdsourcing process. The review contributes to a better understanding of the crowdsourcing process in three ways. First, it proposes a process model to structure the crowdsourcing process. To an extent, the model is more or less aligned with the stages of the crowdsourcing process described in the aforementioned review. Second, the review indicates a strong development of the field, with a significant increase in the number of papers published recently (566 papers). Lastly, Amrollahi (2015), aligning with Zhao and Zhu, concludes the ad hoc feature of the current literature, and further highlights that crowdsourcing process research remains scarce, with only 39 related papers that can be identified out of the 566 papers found.

In summary, the crowdsourcing processes have been studied from both high and low levels of granularity. With high level of granularity, some studies choose an abstract conceptualisation when exploring a crowdsourcing process. As a result, these studies identify main stages and issues that should be considered in the crowdsourcing process. They contribute to the structures of the crowdsourcing process, which enable us to incorporate an analytical framework discussed in the next major section of this review. However, it is important to note that these studies are more focused on highly abstract conceptual understanding and thus detailed activities are still missing.

With low level of granularity, a larger number of studies examine individual processes/events from varied deconstructed aspects. They provide various contributions, reported in case studies, expert opinions, usability studies, experiences, and other engineering development. Though realising the importance of the high-level view, their investigation tends to focus only on parts of the process (Thuan et al., 2017). The ad hoc nature of these studies is repeatedly complained and is highlighted by the two major reviews in the domain. Furthermore, these reviews highlight that research into the crowdsourcing process as a whole is scant, something also suggested by others (Hossain, Kauranen, & Busi, 2015; Mao et al., 2017). As a result, the domain is still unstructured and lacks "a comprehensive guideline through which practitioners can initiate and manage their crowdsourcing projects" (Amrollahi, 2015, p. 2).

To conclude, a few studies cover the crowdsourcing process as a whole without its parts, while a large number of studies investigate the concept through its parts without the whole. The domain is characterised by a large number of ad hoc knowledge sources, which are scattered, varied and sometimes conflicting. This indicates the lack of a solid knowledge base founding the crowdsourcing process. What is also missing is an integrated view of the two levels of granularity, which can provide a complete picture on decomposed activities of the crowdsourcing process and their coordination.

Such an integrated view can be achieved through a business process lens, which has rarely been adopted in the crowdsourcing field. This points us to the concept of business process crowdsourcing, conceptualised in the following section.

2.3.2 Business Process Crowdsourcing

This section explores business process crowdsourcing (BPC). As this concept is relatively new, the review is limited to a small amount of existing relevant literature.

Based on the need for an integrated picture of crowdsourcing processes, this book investigates crowdsourcing using a business process lens. We refer this view as *Business Process Crowdsourcing* (BPC). The term BPC was first coined by La Vecchia and Cisternino (2010) to describe a model allowing organisations to utilise the power of the crowd for their internal business processes, as an alternative to Business Process Outsourcing (BPC). We further define the concept as a way to use crowdsourcing as repeatable organisational business processes. The etymology of the BPC concept is a combination of the phrase *business process* with the word *crowdsourcing* (Thuan et al., 2017). We bring the concept of business process into the concept of crowdsourcing, and consider them as equally important. As the concept of crowdsourcing has been extensive discussed in this book, here we discuss the concept of business process. A business process, according to van der Aalst and Hee (2004), is defined as a combination of individual activities and a workflow describing their logical order. A business process serves as a template for creating multiple, real life instances of the same process, which organisations may create repeatedly and concurrently.

> Given that, this book defines *BPC as a set of activities completed by crowdsourcing entities, in conjunction with a logical coordination of these activities, that collectively form the entire business process.*

Our proposition is that BPC proposes an efficient structured approach for organisations to establish a crowdsourcing process. This efficiency is realised through three roles. First, BPC can help establish repeatable crowdsourcing processes. Inheriting from the business process construct, BPC serves as a template for which organisations create multiple instances of the same repeatable crowdsourcing process. The repeatable characteristic enables analysis of individual aspects of crowdsourcing and their coordination into an organisational workflow (La Vecchia & Cisternino, 2010; Lüttgens et al., 2014). By establishing well-organised workflows, organisations can integrate the crowdsourcing strategy with their day-to-day business processes (Tranquillini et al., 2015). Thus, it enables the incorporation of the crowdsourcing capabilities into the organisational value proposition.

Second, with BPC organisations can start standardising crowdsourcing processes. A pre-condition for process standardisation is that we can comprehend all related activities and their relationships (Thuan et al., 2017). Relying on both the individual and coordinated views, BPC is in a unique position for this comprehension. More precisely, BPC can provide both a detailed view to understand the deconstructed aspects, and a holistic view to understand their relationships, both necessary for process standardisation. This is similar to the role of the business

process view on standardisation of outsourcing (Wüllenweber, Beimborn, Weitzel, & König, 2008).

Finally, BPC contributes to move crowdsourcing toward a more well-defined status. The current ad hoc status of the domain has been noted and discussed in the previous section of the review. Bringing a business process lens to crowdsourcing, BPC allows analysing and defining the basic workflows of crowdsourcing processes, and enabling us to build crowdsourcing processes on top of existing business process management (BPM) technology (Khazankin et al., 2012a; Satzger, Psaier, Schall, & Dustdar, 2011; Tranquillini et al., 2015). In this sense, BPC is expected to efficiently establish crowdsourcing as a common well-defined practice.

Given these important roles, BPC has recently attracted considerable research attention. Many researchers have called for further research on BPC, especially how to conceptualise, establish, and coordinate it. Vukovic et al. (2010) raise "how does crowdsourcing become an extension of the existing business process" (p. 7). Khazankin et al. (2012a) echo similar question and complain about "the lack of an integrated way to execute business processes based on a crowdsourcing [platform]" (p. 1). Similarly, other studies have recently highlighted the demand to build a dedicated crowdsourcing process. This demand increases when organisations have recently used crowdsourcing for core organisational processes like product development (Djelassi & Decoopman, 2013), innovation processes (Lüttgens et al., 2014), industrial processes (Muntés-Mulero et al., 2013), and software development processes (Stol et al., 2017), which have to be coordinated with other organisational business processes.

In spite of these calls, there has been little investigation into BPC and thus how to establish BPC has not been fully examined in the literature. Some prior studies have touched different aspects of BPC. Satzger et al. (2011) seek to help organisations "fully automate[d] deployment of their tasks to a crowd, just as in common business process models" (p. 67), but focus only on choosing suitable workers to perform tasks. Similarly, Khazankin et al. (2012a) highlight the need for organising business processes based on crowdsourcing, but they investigate only a part of the problem, which is how to optimise task properties for supporting business process execution.

A few recent models/frameworks conceptualising crowdsourcing processes contribute to the understanding of BPC. One of the earliest model is proposed by Pedersen et al. (2013). From a process perspective, they in-depth analysed existing research in the domain for conceptualising crowdsourcing. As a result, they propose a conceptual model, organised as an Input-Process-Output structure. The model explains key dimensions of crowdsourcing, including problems, technology, processes, governance, people, and outcomes, which provides a starting point for further study on crowdsourcing processes.

Also analysing existing research in the domain, Hetmank (2013, 2014) aimed at understanding crowdsourcing systems and their components. For this purpose, he suggests a model comprising of four components: user management, task management, contribution management and workflow management (Hetmank, 2013). Based on the identified components, Hetmank (2014) further proposes a lightweight

ontology defining vocabularies of crowdsourcing systems. The vocabularies specify classes and properties, which are useful for crowdsourcing system development. Yet, further evaluation is needed to empirically test the proposition before its practical use, as noted by the author (Hetmank, 2014).

The crowdsourcing process has also been modelled using BPM technology. Tranquillini et al. (2015), based on Business Process Model and Notation (BPMN) technology, modelled workflow patterns of the crowdsourcing processes. They also designed a run-time environment operating these patterns in order to support the workflow enactment. As a result, the study offers a modelling language supporting crowdsourcing workflow enactment and a visual editor that allows organisations to graphically create and manage their crowdsourcing processes. We note that although this work can enact, prototype, and configure a crowdsourcing process, it can only maximise its contribution with the assumption that organisations have already had clear structures of the crowdsourcing process. In other words, this work provides useful supports to configure business process crowdsourcing, which can only be possible if BPC can be clearly established. This further highlights the role of BPC establishment.

Overall, since crowdsourcing needs to evolve from an ad hoc one-off process, we bring the business process lens to research crowdsourcing. We have introduced the concept of BPC and described its possible roles in moving crowdsourcing processes forward. Given these roles, many researchers have suggested further examination of BPC. However, there have been few attempts to do this, and even fewer attempts to establish and support BPC. These attempts have led to a few models/frameworks of crowdsourcing processes, but these models focus primarily on technical features of crowdsourcing systems rather than the business processes orchestrating on these systems. Furthermore, most of the proposed models so far are inconclusive and thus more empirical research is needed (Amrollahi, 2015; Hetmank, 2013). Thus, what is largely missing in the literature is an informed way to establish BPC, from conceptualising, to modelling, and to empirically supporting BPC establishment.

When initially conceptualising BPC, we note that an antecedent must exist to enable the BPC concept. That is, there exists repeatable building blocks of crowdsourcing processes, which provides the process designers basic elements for creating real life instances of the crowdsourcing process. From the preceding review, we have observed several processes, activities, and components that have been repeatedly discussed and thus can possibly be synthesised into the repeatable building blocks of BPC. The following section explores this possibility, leading us to identify the three highly abstract building blocks.

2.3.3 An Analytical Framework of Business Process Crowdsourcing

Investigating BPC, we now present *an analytical framework* decomposing the concept into its abstract building blocks. The framework mainly draws on the existing literature. Such an analytical framework allows us to channel the related research, and later on, will be used to support our analytical process when we analyse a large number of knowledge sources in the domain to identify repeatable business processes of crowdsourcing.

We start with an abstract view on crowdsourcing activities discussed in previous sections, which, by and large, can be grouped into three high abstract stages: decision to crowdsource, design, and configuration. A crowdsourcing process logically starts with a managerial *decision to crowdsource or not*. This managerial decision considers the appropriateness of crowdsourcing to enhance existing organisational tasks (Afuah & Tucci, 2012; Muhdi et al., 2011; Thuan et al., 2013; Thuan, Antunes, & Johnstone, 2016). After the decision to crowdsource, *design* concerns a set of decisions that have to be made to instantiate a concrete crowdsourcing process. We use the term design to highlight the fact that multiple instantiations are possible and that choice depends significantly on subjective criteria. *Configuration* concerns the materialisation of a design into a concrete system (Kittur, Smus, Khamkar, & Kraut, 2011; Little, Chilton, Goldman, & Miller, 2010). These three stages constitute the analytical framework that presents a logical view of the crowdsourcing process. It is graphically presented in Fig. 2.3.

Each stage of the framework and their main concerns are discussed below.

Decision to Crowdsource

The reviewed literature suggests that a crowdsourcing strategy, like other sourcing strategies, begins with go/no-go decision. This decision is referred to as the decision to crowdsource that considers whether crowdsourcing is appropriate for the organisational tasks (Thuan et al., 2016). Muhdi et al. (2011) and Schenk et al. (2017) position the decision to crowdsource in the first-order position starting the crowdsourcing process. A similar position and purpose of the decision is explicitly stated by other researchers (Djelassi & Decoopman, 2013; Lüttgens et al., 2014; Sandkuhl, Smirnov, & Ponomarev, 2016; Wexler, 2011).

The decision to crowdsource plays a central role in a crowdsourcing strategy, for several reasons. First, it is a strategic decision that directly links to whether an organisation will open or close their boundaries to the crowd (Schenk et al., 2017). Second, it affects the use of organisational resources, at least the resources dedicated to crowdsourcing, because inappropriate decisions are likely to lead to unplanned challenges (Rouse, 2010). Furthermore, as a special kind of project that links to the crowd, a failed crowdsourcing project caused by the decision will influence badly on the organisation's reputation (Thuan et al., 2013). Finally, with

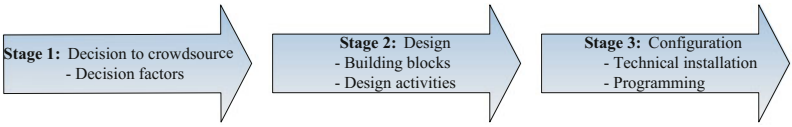

Fig. 2.3 Three-stage analytical framework

its first-order position, the decision to crowdsource cannot be changed, and thus it greatly influences the remaining stages of the entire crowdsourcing processes (Muhdi et al., 2011).

With the first-order position of crowdsourcing, the decision to crowdsource has already received much attention from researchers, focusing on factors driving this decision. Some earlier studies, maybe for simplification, take into account only one factor to make the decision to crowdsource or not. For instance, Ranade and Varshney (2012) addressed the decision "to crowdsource or not to crowdsource" (p. 1) by mainly relying on the factor of task nature. Naroditskiy et al. (2013) examined "the trade-off between the potential for increased productivity with the possibility of being set back by malicious behaviour" (p. 1). However, more recent studies examine a combination of diverse factors in this decision, including benefits and risks (Lu, Hirschheim, & Schwarz, 2015; Muhdi et al., 2011) and organisational structures that founds crowdsourcing operations (Djelassi & Decoopman, 2013). Consequently, the decision to crowdsource is not simple yet complex, where multiple contingency factors should be considered (Afuah & Tucci, 2012; Thuan et al., 2016; Zhao & Zhu, 2014).

Design

After organisations decide to crowdsource, they need to transfer this decision to concrete designs. Design is defined, according to Hevner and Chatterjee (2010), as a plan for structuring elements in order to best accomplish a particular purpose. Adopting this definition, the design stage should plan and structure activities of a crowdsourcing process. It is here the BPC view should maximise its benefits. In other words, this stage should identify both the abstract building blocks, and the detailed design activities and related information structures of the crowdsourcing process.

The literature has shown several possible building blocks of BPC and their detailed design activities. For example, Sect. 2.3.2 has reported a few building blocks proposed by Pedersen et al. (2013) and Hetmank (2013). Another example is the list of design building blocks and activities suggested by Kittur et al. (2013), who research crowdsourcing on complex, large-scale tasks. This list includes twelve abstract activities, including workflow design, task assignment, hierarchy, real-time response, collaboration, quality control, crowds guiding artificial intelligence, artificial intelligence guiding crowds, platforms, task design, reputation, and

motivation. Other design building blocks and their detailed activities can also be found in the literature (Ågerfalk et al., 2015; Amrollahi, 2015; Stol & Fitzgerald, 2014; Zogaj et al., 2014).

Given the existence of different building blocks and their detailed design activities, we note here three important points. First, the differences, again, confirms the ad hoc nature of the domain, and thus suggest a more comprehensive integrated approach to synthesise these building blocks. Second, these studies have highlighted the need to design the different building blocks and activities of the crowdsourcing process. That is, to establish a crowdsourcing process, several activities of the crowdsourcing process need to be designed and structured, which suggests the role of process design (Stage 2 of the framework). Finally, despite the differences of the proposed building blocks and activities, we can identify some repeatable activities, such as crowd management (Kittur et al., 2013; Pedersen et al., 2013), how to motivate the crowd (LaToza & Hoek, 2016; Naderi, 2018), and quality control (Amrollahi, 2015; Kittur et al., 2013). Consequently, it is possible and necessary to reconcile the differences and suggest common building blocks of how to design the crowdsourcing process.

Configuration

The configuration stage transforms a crowdsourcing design into a concrete implemented system. In the crowdsourcing context, configuration can refer to either technical decisions to set up crowdsourcing components on existing platforms (Gonnokami, Morishima, & Kitagawa, 2013; Hosseini, Phalp, Taylor, & Ali, 2015a; Kittur et al., 2011; Little et al., 2010), or in-depth technical software development to build a crowdsourcing platform, such as algorithms, protocols, and database structures (Schall, 2012). Although this stage can be considered from both views, the chosen business process perspective limits our concern within the process configuration on an existing platform. This is also supported by the availability of several crowdsourcing platforms (Hirth et al., 2011) and programming toolkits that eases the configuration (Kittur et al., 2011; Kucherbaev et al., 2013; Little et al., 2010; Tranquillini et al., 2015).

Overall, we have synthesised the analytical framework initially conceptualising BPC. The framework structures three high-level stages of BPC: decision to crowdsource, design, and configuration, which will be deconstructed into the main building blocks and activities to thoroughly conceive the BPC concept.

2.3.4 Discussion

The review assessed the literature on crowdsourcing processes and business process crowdsourcing. It identified the two major research streams of crowdsourcing processes: high and low levels of granularity. It finds that some studies research crowdsourcing processes as a whole without its parts, while a large number of studies investigate specific parts of crowdsourcing processes without the whole.

The different levels of granularity have hindered us to have a completed picture of crowdsourcing processes. Further, major reviews of crowdsourcing literature show that the domain is characterised by a large number of scattered, varied and sometimes conflicting knowledge sources. Addressing this challenge requires an integrated view, which has led us to introducing BPC.

Reviewing what little has been published on BPC highlights three important points. First, the review introduces the concept of BPC that brings a business process lens to study crowdsourcing, which enables us to establish crowdsourcing as an organisational business process. Second, the review discusses the roles of BPC. It shows that BPC can resolve the ad hoc challenge and provide structures for the domain. Finally, it finds a few models and frameworks contributing to understand crowdsourcing processes, but not comprehending BPC. Together, these points suggest that BPC is an emerging yet important phenomenon that needs to be conceptualised, modelled, and applied to crowdsourcing practices.

Despite of its early state, BPC is promising to move crowdsourcing from ad hoc processes toward mature repeatable processes. That is, BPC provides a template of repeatable building blocks that organisations can use to instantiate real-life crowdsourcing processes. From the preceding review, we observe that some building blocks that have been repeatedly discussed. Moving this observation forward, we initially synthesise three abstract stages of BPC repeatedly suggested in the crowdsourcing process literature. These stages allow us to channel the related literature in the next chapters to obtain increased insight and thoroughly conceiving BPC.

We note that from the current early state of BPC, this book will engage in conceptualising, modelling, and supporting business process crowdsourcing. The resulting engagement, presented in the remaining chapters of this book, contributes to move the domain to a more mature state, which will be further discussed in Sect. 7.2.4.

2.4 Chapter Summary

This chapter has provided a narrative review to assess the state of the art that driving the book to study BPC. One main drive is that the crowdsourcing domain is emerging, characterised by unstructured knowledge sources and the lack of a strong knowledge base. There appears to be evidence for this in the literature reviewed in the previous sections.

The review covered three major strands. The first strand examined the conceptualisation of crowdsourcing. It shows three main pillars behind the concept, followed by a discussion in order to compare and contrast crowdsourcing with other related concepts. They draw a boundary around the crowdsourcing concept and show that crowdsourcing is a distinctive concept per se. Then the short history of crowdsourcing definitions was discussed to show that the concept continues to evolve. Together, the distinctive concept suggests that crowdsourcing must be

developed independently, while the evolution of the concept's definitions and its short history indicate the emerging nature of crowdsourcing.

The second strand reviewed basic classifications in the crowdsourcing domain. This shows that research into classifications cover many particular topics: applications, tasks, types of crowd members and platforms, but not yet cover the synthesis and coordination among them. It is the ad hoc nature of the domain, where classifications are solely relevant to particular crowdsourcing elements or contexts. Further, these classifications have not yet been synthesised and linked in a comprehensive integrated structure, and thus there is still a need for a solid knowledge base that structures the domain.

Finally, the last strand has painted an overall picture of business process crowdsourcing. It shows that BPC is still in an early state with a small amount of related literature, which needs to be further conceptualised, structured, and supported. At the same time, the review shows that BPC is important to establish repeatable crowdsourcing processes, and thus possibly moves the domain toward more mature state. To contribute to this movement, the review has developed an analytical framework presenting the three abstract stages of BPC: decision to crowdsource, design, and configuration. The framework abstractly conceptualises BPC, and will be used to guide our data collection and analysis for further conceptualising and structuring BPC.

The following chapter discusses the main building blocks of BPC.