CROWDSOURCING DESIGN REFERENCE: A PRELIMINARY MODEL

Nguyen Hoang Thuan, SIM, VUW.
Thuan.nguyen@vuw.ac.nz

1 AProf Pedro Antunes
2 Dr David Johnstone

Abstract
How to integrate crowdsourcing with organizational business processes is still an open question. Addressing this question, the current study analyses the integration as a sequence of stages, including decision to crowdsource, design and implementation of a crowdsourcing system. Extending our previous work investigating the first stage, this study conducted a structured literature review focusing on the other stages: the design and the implementation. As a result, a preliminary reference model guiding each stage in the crowdsourcing integration is proposed.

Keywords
Business process, crowdsourcing, component, design, implementation, preliminary model

Introduction
The dominance of Web 2.0 has changed the ways of online interaction, in which the members of the crowd have evolved from passive receivers to active contributors (Brabham, 2010). Crowdsourcing, an emerging model utilizing the ability of these active members in the crowd, is promising to bring competitive advantages for organizations. Pioneering studies have suggested that organizations by choosing a crowdsourcing strategy can harvest expertise, innovation, information, skills, and labour from the crowd (Aitamurto, Leiponen, & Tee, 2011; Brabham, 2008; Vukovic & Bartolini, 2010), and thus achieve flexibility, responsiveness, and low cost for their business strategies (Hetmank, 2013; Whitla, 2009).

With these advantages, crowdsourcing is expected to widely integrate with organizational business processes. However, this is not likely to have happened, as seen via the fact that only 10% organizations have actually deployed a crowdsourcing strategy (Andriole, 2010). One possible explanation for this low rate is that deploying a crowdsourcing strategy is risky (Hetmank, 2013) since “[organizations] do not know how” (Malone, Laubacher, & Dellarocas, 2010). Therefore, researchers are advocating to further investigate the question, how to integrate crowdsourcing with organizational business processes (R. Khazankin, Satzger, & Dustdar, 2012; Vukovic & Bartolini, 2010).

This study decomposes the integration into a sequence of stages involving the decision to choose crowdsourcing, the design and the implementation of the crowdsourcing system. Using a divide and conquer strategy, such decomposition eases answering the aforementioned question. Previously, we conducted an exploratory study on the first stage of the integration process and identified the factors related to the decision to crowdsource (Thuan, Antunes, & Johnstone, 2013). The current study concentrates on the other stages, exploring how to design and implement the corresponding crowdsourcing system. To address this question, interdependent components comprising a crowdsourcing system need to be clarified before the design and implementation stages. However, the literature (i.e. D Geiger, Rosemann, Fielt, & Schader, 2012; Hetmank, 2013; Pedersen et al., 2013) is currently inconsistent on differently proposing the components of the crowdsourcing system. Thus, further study is necessary to synthesize these components.

Responding to this necessity, this paper extends the analysis of the structured literature review conducted in our previous work (Thuan et al., 2013), in order to identify the components that should
be considered in designing and implementing a crowdsourcing system. The study contributes to
knowledge by proposing a reference model of components, which are important in designing and
implementing crowdsourcing system (Hetmank, 2013). As a result, it adds a further step to answer the
question, how to integrate crowdsourcing with organizational business processes. In practice, the
model can help organizations have a complete picture on planning and deploying their chosen
crowdsourcing systems.

Literature Review

Since identifying components of a crowdsourcing system is the initial but key phase before these
components can be described and developed, several studies have proposed models and frameworks
of these components. Kittur et al. (2013) focusing on complex and large scale crowdsourcing tasks
propose a map of twelve foci, which summaries “issues of design – the technical and organizational
mechanisms surrounding crowd work” (Kittur et al., 2013). There are two reasons that Kittur et al’s
work is related to the current study. First, on each of their proposed foci, the discussion involves
important research issues and themes, which serve as the starting point for identifying components in
a crowdsourcing system. Second, although this work does not specify the crowdsourcing integration
with organizational business processes, it using the theory of organizational behaviour to explain the
relationships between crowdsourcing with organizations, individuals and technologies provides a
frame for examining the integration.

Using structured literature review methods, other researchers aim to articulate current findings from
the crowdsourcing studies for identifying the components. Pedersen et al. (2013) chose three IS
conferences and 11 IS journals to survey crowdsourcing studies. Based on classical Input-Process-
Output model, they present a conceptual model of corresponding components, which label as
problem, outcome and four components between the problem and the outcome including technology,
process, governance, and people. However, some limitations can be found in this study. One of them
is that the study constrains to IS literature while other research areas, such as business, include many
valuable crowdsourcing studies (i.e. Howe, 2006; Whitla, 2009). Second, the paper does not include
studies on crowdsourcing process, as seen via “we did not find research specifically related to
understanding the process used in crowdsourcing applications” (Pedersen et al., 2013), but in fact this
topic has examined by many crowdsourcing studies, as summarised in the taxonomy of
crowdsourcing processes by Geiger et al. (2011). Finally, the proposed conceptual model does not
explain the relationships between its components and how to integrate these components to the
organizational context.

Without restricting to IS studies, Hetmank (2013) recently reviewed 220 research papers, which were
extracted from eight online bibliographic databases. The author suggests other components of a
crowdsourcing system, such as user management, task management, contribution management and
workflow management. Within each component, its related functions are recommended. Although
this study is quite comprehensive in term of scope of the surveyed literature, it has a quite narrow
focus due to its chosen technical perspective. As a result of this narrow focus, the study only
examines the technical integration between crowdsourcing and IT environment, and thus the link
between crowdsourcing and the business processes operating on this IT environment is still missing.

In summary, the recent studies have highlighted the importance of identifying components in a
crowdsourcing system, which are necessary for the integrating crowdsourcing with organizational
business processes. However, the aforementioned studies chose either too abstract (e.g. Pedersen et
al., 2013) or too technical perspective (e.g. Hetmank, 2013), leading to two consequences. First, the
integration between crowdsourcing with existing business processes is not comprehensively addressed
although some studies (i.e. Hetmank, 2013) do investigate the technical integration. Second, different
studies have suggested different “key components” that should be considered in the design and
implementation of the crowdsourcing system. Taking that in consideration, this study aims to fulfil
these gaps by proposing a reference model guiding the design and implementation of the
crowdsourcing system, leading to the integration of crowdsourcing with existing business processes.
Research method

This study chose the structured literature review method, which is similar to our previous work used to review the decision to crowdsource (Thuan et al., 2013). In the previous work, a search was conducted on six online bibliographic databases. Using ‘crowdsourcing’ as the keyword, the search was performed between February and March 2013. As a result, 611 publications were found (Table 1). By eliminating editorial introductions, duplicates, posters, letters, and tutorials, the initial pool includes 500 papers.

<table>
<thead>
<tr>
<th>Document types</th>
<th>ACM</th>
<th>IEEE</th>
<th>Science Direct</th>
<th>Sage</th>
<th>Emerald</th>
<th>Springer Link</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conference paper</td>
<td>274</td>
<td>110</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>384</td>
</tr>
<tr>
<td>Journal</td>
<td>-</td>
<td>33</td>
<td>33</td>
<td>16</td>
<td>8</td>
<td>137</td>
<td>227</td>
</tr>
<tr>
<td>Total</td>
<td>274</td>
<td>143</td>
<td>33</td>
<td>16</td>
<td>8</td>
<td>137</td>
<td>611</td>
</tr>
</tbody>
</table>

*Table 1. Crowdsourcing publications, search result (Thuan et al., 2013)*

Focusing on crowdsourcing design and implementation publication, the current study examined the initial pool and chose the related papers. Based on the aforementioned foci proposed by Kittur et al. (2013), this study inspected the papers’ titles and their keywords. As a result, 112 articles related to the design and implementation phase were classified. Although the final target is to analyse all these articles, the current study randomly chose 40 articles for its pilot analysis due to the available amount of time.

To identify the components from these chosen articles, this study adopted a combination of deductive and inductive approach for its analysis. On the one hand, some themes, which were presented from the literature review section and from previous works (Kittur et al., 2013) need to be taken into account in a deductive approach. On the other hand, this study aims to explore the frequent and dominant components from the surveyed literature without limiting by the predefined themes, and thus an inductive approach is suitable (Thomas, 2006). Taken together, the combination of deductive and inductive approach was chosen for this study.

A preliminary model

By analysing the chosen articles, six components: task definition, workflow design, crowd management, quality control, incentive mechanism and implementation, and three related elements: types of tasks, the crowd, and types of platform can be synthesized. Combining them with the decision to crowdsource, identified by our previous study (Thuan et al., 2013), a preliminary reference model is proposed in Figure 1.
The initial component in the model is the decision to crowdsource. Although this component is excluded from the crowdsourcing process by some studies (i.e. Zhao & Zhu, 2012), the reference model follows suggestions from others (i.e. Afuah & Tucci, 2012; Sutherland, 2013; Wexler, 2011) to include this component. Our previous work (Thuan et al., 2013) examined this component and identified the factors related to the decision to crowdsource, including task properties (internet vs. physical, interactive vs. independent, sensitive information vs. non-sensitive, partitioned vs. non-partitioned), people factors (employee and the available of the crowd), management factors (budget, experts and risk), and environment factor (the available of crowdsourcing platform).

Task definition is presented as the second component in the model since the existing literature highlights the role of tasks in crowdsourcing project and suggests clearly defining tasks being crowdsourced (Malone et al., 2010; Rosen, 2011). A crowdsourcing task can be initially defined through classifying it into proposed types of tasks. Although different ways to structure types of tasks are suggested (Brabham, 2012; Whitla, 2009), the current study chooses the way proposed by Schenk and Guittard (2011), with two dimensions: task complexity and the nature of task. Task complexity includes three levels simple, complex and creative. The nature of tasks reflects how tasks can be performed: individually or competitively (Schenk & Guittard, 2011). Although these two dimensions are simple, they help to understand what crowd should be approached and how tasks can be performed, which are essential in task definition.

After identifying the type of task, other task’s properties should be specified in task definition. Zheng et al. (2011) focusing on crowdsourcing contests suggest to clarify four properties, including autonomy, variety, tacitness, and complexity. Choosing a broader approach without limiting to crowdsourcing contests, Tokarchuk et al. (2012) go further in recommending five properties: autonomy, variety, identity, specificity, and significance. According to Tokarchuk et al. (2012), autonomy property defines how much freedom regarding time and quality a member has to follow when performing tasks. A task with high variety requires members having skills and performing a variety of activities to accomplish it. To define identify and significance of a task, organizations need to show when finishing the task how its result contributes to the final outcome and the outside world respectively. Finally, specificity property refers to the requirements of the task, which is low if the task is very open and high if the task is highly specific. In short, these properties should be specified in defining crowdsourcing tasks, and depending on a particular task other properties may be suggested.

Third, workflow design is a fundamental component in crowdsourcing system (Kittur et al., 2013; Kulkarni, Can, & Hartmann, 2012). Two research streams related to the workflow design are task decomposition and result aggregation (Kittur, Smus, Khamkar, & Kraut, 2011). Although most of the current studies in these streams address simple tasks, task decomposition and result aggregation may also be applied for other kinds of tasks like problem solving. According to Afuah and Tucci (2012), decomposing these tasks “may be easier for the focal agent to articulate a module”. As a result, workflow design is suggested as an important component in crowdsourcing system for not only simple but also complex and interdependent tasks (Kittur et al., 2013).

In crowdsourcing projects, tasks are accomplished by the members in the crowd. Thus, the crowd’s characteristics also attract a large number of studies. Considering crowd’s characteristics, it is important to note that not all the member in the crowd will perform tasks, and some do perform but do not finish the chosen tasks. According the SCOUT framework (Stewart, Lubensky, & Huerta, 2010), which separates the members in the crowd to super contributors, contributors and outliers, there are about 67% members in the crowd contributing efforts to accomplish tasks, and within these members only 1% give super efforts. Recently, Chandler and Kapelner (2013) conducted an experiment in a crowdsourcing platform supporting micro tasks. They find that within the members participating to accomplish tasks, 25% begin but not complete their chosen tasks. From these observations, how to make sure that the crowd’s members accomplish crowdsourcing tasks is challenge.

A widely chosen approach to address this challenge is to profile the crowd and assign tasks. One of the key actors in this direction is the research group in Technical University of Vienna, who suggest using the members’ interests and skills to build the members’ profiles that will be updated based on
the accomplished task quality (B Satzger, Psaier, Schall, & Dustdar, 2011). Based on these profiles, this group suggests different mechanisms to assign tasks to suitable members, such as auction based mechanism (B Satzger et al., 2011; B. Satzger, Psaier, Schall, & Dustdar, 2012) and scheduled crowdsourcing, allowing individuals in the crowd to negotiate the quality and deadline to finish tasks (Roman Khazankin, Schall, & Dustdar, 2012). Besides this group of researchers, profiling the crowd and assigning tasks also attract other researchers (i.e. Kern, Thies, Bauer, & Satzger, 2010; Kucherbaev et al., 2013; Venetis, Garcia-Molina, Huang, & Polyzotis, 2012).

The fifth set of studies has concentrated on investigating different strategies for quality control. Since crowdsourcing is based on voluntary workers who are members of a large and anonymous crowd (Hirth, Hoßfeld, & Tran-Gia, 2012), the fact that no formal commitment exists between workers and requesters may lead to quality issues. Zhao and Zhu (2012), in a recent literature review, summarize three main strategies that can be used for quality control: experts, the crowd, and third-party organizations. Other studies have examined each strategy to clarify its advantages and disadvantages, and provide guidance for selecting the suitable one in the organizational context (Hirth et al., 2012; Kern et al., 2010; Vukovic & Bartolini, 2010).

The success of a crowdsourcing system is directly influenced by a large number of crowd’s members, and thus how to motivate members to participate in a crowdsourcing system or incentive mechanism has been studied by many studies. The motivation can be divided to intrinsic motivation and extrinsic motivation (Brabham, 2010; Kaufmann, Schulze, & Veit, 2011). Whereas Kaufmann et al. (2011) focusing on micro tasks find that intrinsic motivation is more important compared to extrinsic one, Brabham (2010) examining a competitive platform argues that extrinsic motivation is more significant than intrinsic one. Although further study is needed to clarify this inconsistence, the current study believes that both motivations need to be considered in the design of the crowdsourcing system.

Lastly, research also focuses on how to implement crowdsourcing in a certain platform. Some studies in this group investigate how the crowdsourcing processes can be programmed in a specific type of platform (Ahmad, Battle, Malkani, & Kamvar, 2011). Others concentrate on governance the crowd in order to receive qualified and scheduled results (Bozzon, Brambilla, Ceri, & Mauri, 2013). By allowing the crowd to participate, organizations may lose their control and focus (Zhao & Zhu, 2012). This is where governance mechanism to control the crowd is necessary. For instance, Jain (2007) examines three crowdsourcing initiatives and presents different governance mechanisms that can be used in the implementation phase.

**Conclusion and Limitations**

In summary, this study using a structured literature review method proposed a preliminary reference model of seven components that should be considered in the design and implementation of a crowdsourcing system. Considering both technical and managerial components, the reference model can be seen as an extension of the work by Hetman (2013), who only focuses on technically integrating crowdsourcing with IT environment. As a result, the model provides an overall picture to guide the integration of crowdsourcing with not only the IT environment in the organization but also organizational business processes.

When critically reflecting this work, some potential improvements can be applied. First, the model is proposed based on the analysis of 40 articles, which should be extended in the future research. Second, the crowdsourcing integration was decomposed into three stages in this study. Although such decomposition helps to address the research problem, clearly defining the line between these stages is difficult and thus some components in one stage seem to fit in the other stage. For instance, ones can argue that task definition should not be embedded in the design stage, but in the decision to crowdsour. However, the current study believes that although some task properties should be examined in the crowdsourcing decision, as seen via our previous work (Thuan et al., 2013), defining task in detail is crucial in the crowdsourcing design phase.
References


