Choosing the Right Crowd: An Iterative Process for Crowd Specification in Crowdsourcing Initiatives

Eoin Cullina
LERO, NUI Galway
e.cullina2@nuigalway.ie

Kieran Conboy
LERO, NUI Galway
kieran.conboy@nuigalway.ie

Lorraine Morgan
LERO, NUI Maynooth
lorraine.morgan@nuim.ie

Abstract

Crowdsourcing enables organizations to obtain dynamic solutions from large and diverse crowds of individuals. With the advent of IT-mediated ecosystems crowdsourcing initiatives readily possess a means to mediate, regulate and refine crowdsourcing sub processes including the process of vetting and selecting crowd membership (crowd specification sub process). Despite this fact there is ambiguity in current research as to what constitutes a crowd. Furthermore, traditional crowd selection processes show signs of limited application and are often conducted (i) once, at a fixed stage in the process and (ii) by the process initiators. This research-in-progress paper seeks to address constraints caused by the limited application of these processes. Firstly, a definition of the crowd is formulated from concepts identified in literature. Secondly, an iterative conceptual model is advanced from theory that facilitates the creation of a crowd tailored for tasks of varying size and complexity.

1. Introduction

The emergence of open innovation [7] and the subsequent concept of crowdsourcing [26] challenges the traditional viewpoint that the best solutions always come from experts inside an organization. Historically the leading examples of innovation in industry came from small internal projects [23]. From the 1960s through to the present the behavioral sciences advanced research into collective behavior, herd behavior and crowds [53, 39, 24]. Recently attempts have been made to define emergent phenomena such as crowdsourcing [13, 49, 17, 20]. Numerous advantages have been identified in the process over more traditional solution sourcing processes. Studies show that crowdsourcing initiatives amongst users can under certain conditions outperform professional in-house activities for the generation of new product ideas [43]. Research has also shown that in certain instances that the best solutions presented by crowd members come from outside the presenting field or domain [29]. Lakhani asserts that experts in fields outside of the problem domain can have access to a wider range of methods and design paradigms which they may be able to apply to the problem in unique and innovative ways [15, 29]. Crowdsourcing as a concept has changed and evolved over the years. It started life as a commercial process for application in business environments but has since been used by other types of non-commercial actors such as governments and charities [6]. Crowdsourcing has been used for a kaleidoscope of activity ranging from the completion of menial tasks (micro-tasks) [31] to the resolution of complex scientific challenges through gamification [22]. Moreover, platforms such as Amazon Mechanical Turk (AMT) and InnoCentive have accelerated the success of crowdsourcing through their usage by the masses. What sets crowdsourcing aside from other traditional solution solving processes is the concept of the crowd as a dynamic agent of change. Yet despite this apparent dynamism aspects of the process remain constricted and limited. Many of these problems and restrictions can be traced back to uncertainty surrounding core concepts at the center of crowdsourcing. This paper seeks to focus on one specific sub process within crowdsourcing namely the crowd specification process. This paper adopts a unique approach by introducing iterative cycles to the crowd specification process. The use of such iterative refinement cycles is evident in areas such as agile software development.

2. Motivation and research objectives

At a fundamental level it is unclear as to what constitutes a crowd within crowdsourcing. Some attempts have been made to define crowdsourcing [13] but little agreement has been reached as to what constitutes a crowd. Various definitions of crowdsourcing highlight concepts central to the crowdsourcing concept. In the case of this specific research, matters are further complicated by the fact that in many crowdsourcing initiatives the crowd specification process is often conducted at a fixed stage

early on in the process and is static in nature. This generally assumes one of two forms. In the first scenario, the process initiator provides an unrestricted open call to a large and undefined crowd. In the second scenario the process initiator places clear restrictions upon crowd membership criteria from the outset. Such actions can serve to limit the effect of the crowdsourcing initiative from the outset for several reasons. Firstly, in many cases the initiators of the crowdsourcing challenge are not always best placed to select a crowd matching the exigencies of the task at hand. Secondly, crowd membership cannot be tailored to meet the changing needs of many projects where tasks and challenges vary in terms of size and complexity over time. Accordingly, a new dynamic means of crowd specification is required. It is posited that a one-size-fits-all approach to crowd specification does not suit every type of initiative. Much of the present crowdsourcing research from the information systems (IS) perspective has focused upon the effect of incentivization on process outcomes [25, 60, 10]. To the contrary this research is focused upon achieving greater dynamism in the crowd specification sub process. To this end the problem is rooted within the IS paradigm lying at the intersection of people, processes and technology. IS research is frequently criticized in lacking relevance to practice [47, 32, 5, 16]. Also it is argued that IS research has not fully engaged with crowdsourcing as a phenomenon [38]. This research seeks to contribute to the IS body of knowledge through addressing the challenges above by advancing a definition and a conceptual model. To address the specific issues outlined above the following research objectives are presented:

O1. Identify the core criteria through which crowd membership is assessed.

O2. Present an iterative and non-restrictive crowd selection sub process.

To achieve the goal sought in objective one (O1) a definition of the crowd is consolidated from constructs identified in the literature. In addressing objective two (O2) the crowd specification sub process is forged from its constituent constructs and is presented in the form of a dynamic conceptual model that will be further validated in future stages of this research. The next section of the paper outlines the structure of the literature review completed. The subsequent sections of the paper are presented as follows: Section 4 identifies the core characteristics of a crowd engaged in crowdsourcing and advances a definition from these characteristics. Section 5 addresses size and complexity as issues within crowdsourcing initiatives. Section 6 outlines the crowd specification sub process components. Section 7 discusses cyclical iterative crowd processes. Section 8 presents the next proposed stages of the research. Section 9 deals with the research limitations. Lastly, section 10 presents the research conclusions and potential future areas of research are discussed.

3. Literature Review

3.1 Step 1: Literature review approach

Through categorizing constructs into themes within the literature review process a definition and conceptual model are advanced. The conceptual model is a representation of theory as to the constructs identified and their inter-relationships. Okoli and Schabram state that a literature review "*must contribute to the work in its dual approach of synthesizing the available material and offering a scholarly critique of theory*" [42]. Webster and Watson confirmed the importance of the literature review within IS in describing same as a "foundation" [59]. Levy and Ellis further expanded the literature review technique by adopting a systematic approach to the literature review process [35]. More recent examples of the use of the literature review process advocate the adherence to a strongly systematic method [30]. The authors through the literature review process address both IS literature and research across disciplines regarding crowdsourcing and related phenomena. Concepts are grouped into themes through a systematic approach. The literature research comprises three tiers. 1. A general sweeping review of crowdsourcing literature across disciplines. 2. The focused identification of thematic concepts relating to crowd specification. 3 The synthesis of concepts into constructs for the purposes of forming a definition of the crowd and the construction of a model.

3.2. Step 2: General database searches

In starting with a broad review materials examined included peer-reviewed journals, conference proceedings, web articles and web portal content. Research areas across disciplines were examined including labor markets, crowdfunding, citizen science, collective intelligence, collaborative innovation, outsourcing, open innovation and open source software development. A multitude of other related areas also show examples of processes similar to crowd specification within crowdsourcing including inner source and collaborative budgeting. Databases queried included EBSCO Business Source Premier, Emerald

Management Xtra, Elsevier ScienceDirect and Google Scholar. Keyword search terms and associated results included [crowdsourcing], [crowdfunding], [open innovation], [citizen science], [peer production], [collaborative design], [collaborative budgeting] and [wisdom of crowds]. Following the searches a broad initial list of results was compiled.

3.3. Step 3: Focused searches

Subsequent focused searches were completed in the identification of thematic groupings from the broad list of results obtained previously. Peer-reviewed journals containing empirical studies were primarily sourced. Papers were included where clear research methods were discernable. Furthermore, research was prioritized where the study was on crowdsourcing rather than just the use of crowdsourcing. Several core challenges were faced in tailoring the structure of the literature review method at this juncture. Many literature review processes feature the exclusive use of peer-reviewed journals. Many new and novel features included in crowdsourcing platforms have not been fully addressed in the peer-reviewed literature to date. Crowdsourcing is a relatively new phenomenon. The literature review method accordingly has been extended to include web materials and content provided by crowdsourcing platform/portal providers. Furthermore, many open innovation tournaments discussed in literature are occasionally identified as crowdsourcing initiatives. Whereas the terms used are different the concepts contain the same meaning. A part of the review processes included the identification of differing terms holding the same meanings.

Exclusion criteria included initiatives that were entirely focused upon the crowdsourcing of data with no discussion as to the composition of a crowd or the crowd specification process. Publications that exclusively focused upon participant motivations without discussion of the crowdsourcing process were also removed. Through the initial search process some 404 publications of merit were identified and reviewed in a broad capacity. The references cited within these papers were examined for cross-linking with other publications. Through this process some 60 publications were selected as been of particular relevance to the construction of crowd specification sub processes where (i) crowd attributes were presented (ii) process elements were identified or discussed and (iii) relationships between constructs can be identified. In the next section we present a definition for the crowd based upon core components identified in the literature.

4. Defining the crowd

The definition of what constitutes a crowd has somewhat been lost in the attempt to define the crowdsourcing process. Anastasiou & Gupta make an interesting insight into the definition of crowdsourcing stating that it is a 'trend' [2]. From the definitions of crowdsourcing advanced in literature we can see confusion as to the process components and the definition of the crowd [13]. The following thematic groupings have been identified from the literature for the purposes of defining the crowd:

Crowd Size: There appears to be no specific agreed delineation as to what constitutes the size of a crowd. Jeff Howe presented the term as an undefined (and generally large) group of people. It is deduced, as a plural noun, that at its minimum the crowd must feature two or more entities in its membership. Other definitions presented confirm the view that the crowd is seen as generally or usually a large group of people.

Human Crowd: Some authors have placed a restriction upon the crowd that it is exclusively human in composition [11]. However, examples exist with portals such as Wikipedia whereby bots supplement the crowdsourcing process through the performance of tasks otherwise completed by a human crowd [21]. Therefore, we posit that the potential exists in the future for a crowdsourcing process to feature no human contribution and retain bots or artificial agents for the completion of certain tasks. Science fiction and research has suggested that artificial intelligence might be attained in the future.

Known versus Unknown: In present definitions advanced we see further examples of broad and narrow interpretations of what constitutes a crowd. At one end of the spectrum the crowd members are generally unknown. Schenk and Guittard state "*a crowd can be defined as a large set of anonymous individuals*" [50]. This example runs contrary to examples of internal crowdsourcing initiatives used by large organizations [61] and initiatives where certain crowd members can be identified by their skillsets advertised.

Internal versus External: Various examples exist of internal and external examples of crowdsourcing [52]. Although the vast majority of examples include an individual or organization calling upon an external crowd examples exist of government agencies using internal crowdsourcing processes to source new ideas or solutions to challenges [28].

Types of Participation: Evidence exists of crowds interacting in platforms in an active capacity and a passive capacity. Active participants engage in challenges and display behaviors such as voting, liking, competing and collaborating within the system. Passive participants can support an initiative by their mere

presence on the platform. In some cases they may not vote in competitions. However, they may play an invaluable role by assisting other crowd members in their participation. The mere presence of passive participants on the platform is sufficient to be classed as a form of engagement.

Varying Skill: One of the key challenges facing crowdsourcing projects is the matching of crowd members to tasks based upon human skill [46]. Some portals have truly embraced the concept of engaging a broad and dynamic crowd. With InnoCentive, Fried notes that "*to begin browsing and working on unsolved problems, a solver need only create an account on InnoCentive's site*" [15]. Such are the low level of restrictions placed upon membership through InnoCentive. In its broadest form crowdsourcing initiatives place no restrictions on crowd membership based upon skill. Despite this fact examples exist on AMT where only crowd members of a particular skillset can participate in the completion of certain tasks. Crowdsourcing tasks have been classified by Schenk and Guittard (2011) into; routine, complex, or creative tasks [50]. Different types of task require the use of different skillsets from the crowd member. Some tasks can be solved by any individual, whereas others may require additional skills or expert knowledge [8].

Outside of the minimum criteria required to define the crowd, several other factors of note are consideration. Firstly, in many cases the process initiator is treated as an external agent to the crowd. In many examples the *process initiator* engages the crowd through a portal in a tournament and selects a winning idea. However, other examples exist of crowdsourcing initiatives where the initiator organization forms an active part of the crowd in participating and voting upon ideas [27]. Vukovic defines the initiator as a "*crowdsourcing requestor*" namely "an entity that submits a task request – initiating the process of crowdsourcing, by specifying the acceptance criteria [56]. It is not correct in all instances to present the process initiator as external to the process. IBM Innovation Jams have featured participation from both employees of the company and external actors and stakeholders in the same crowdsourcing process.

Secondly, although it is presumed in many instances that the crowd interact through the Internet examples exist of crowdsourcing using mechanisms other than through *web 2.0 technologies* [18]. Suggestion drop boxes and SMS messages over GSM can all be compiled within crowdsourcing initiatives. The fact that the submission is not received in a web 2.0 format does not prevent the process from being labeled crowdsourcing. Although pertinent to understanding how the crowd is formed neither of these points are relevant to the definition of the crowd and accordingly are not addressed in the model of the sub process presented. In light of these facts a revised reflection upon crowd composition is proposed. Accordingly, within this spectrum of diversity we synthesize a definition that encompasses all types of crowd so as to be overarching in nature:

"*Within the crowdsourcing process a crowd is comprised of two or more solving agents who interact through a platform in an effort to address or solve a challenge. Crowd members can be known or unknown, internal or external to the initiator organization, display varying levels of skill and varying types of participation within the process*".

5. Size and complexity

Within crowdsourcing initiatives we see two predominant areas of focus within challenges namely size and complexity. Crowdsourcing can be used to address tasks where the tasks requiring completion are too complex for completion by computers. Furthermore, in some cases large volumes of participants are required for the completion of large-scale tasks. A broad range is evident in the types of tasks addressed in crowdsourcing initiatives. Firstly, where the task at issue requires a level of skill not capable of completion by a computer, crowdsourcing can supplement existing IT processes within the organization. In doing so multiple tasks can be completed in parallel (ReCAPTCHA is one such example) [34]. Secondly, crowdsourcing can be deployed where a task number or task volume count is too excessive for traditional systems to address. The search for flight MH370 through vast amounts of satellite photos presents such an example [57].

Different processes can be used to manage task size and complexity. Many examples exist in practice of the process of breaking down problems into digestible chunks. This process as part of a cycle is known by different names such as task disaggregation and selecting task granularity. The activities of task disaggregation / selecting granularity have a direct effect on the potential target size of the crowd broadcast reach. With ReCAPTCHA two challenges are completed through the process. At one level the user through entering the captcha confirms to the system that the user is human. At the other end of the transaction the user converts handwritten texts into typeface. The level of task disaggregation has a direct effect upon the complexity of the task to be completed

6. Crowd specification sub process

In constructing a crowd specification sub process from theory several core steps are taken. The conceptual model presented is a 'design and action' form of theory [19]. The process initiator or crowd member can start the process of selecting a suitable crowd to complete a specific task. The prospective crowd members are vetted and selected for participation. Recent research has outlined key considerations that shape crowd construction [45]. These include identifying the people required to perform the task and understanding their motivations for participation. Also simple rules need to be extended for the purposes of constructing the crowd [45] pp 77.

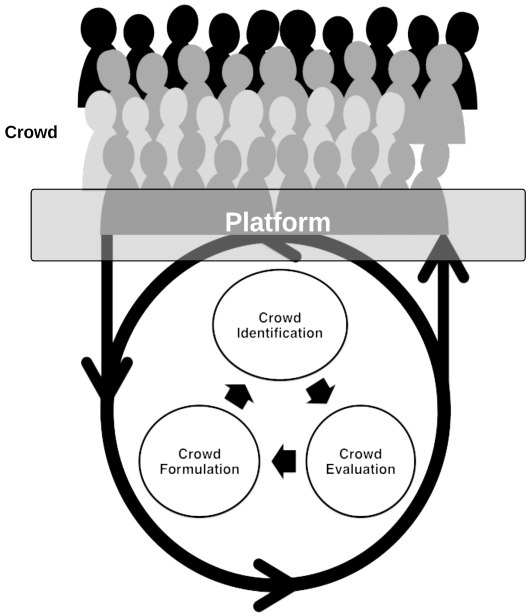

Figure 1. Crowd Specification Sub Process

It is posited that the crowd specification process can be improved through the adoption of iterative cycles within the process. Within the process we see numerous concepts underpinning the criteria by which crowds are built. These concepts can be grouped at a higher level of abstraction into three constructs that form part of the crowd specification process namely *crowd identification*, *crowd evaluation* and *crowd formulation*. These three stages identified encompass all the lower level activities including selecting crowd member skillsets, defining internality / externality etc. The start of the crowd specification process is preceded by the communication of an "open call". The meaning attributed to the term "open call" can be somewhat ambiguous. The call itself can be subject to quite a number of restrictions. This does not restrict the process initiator from broadcasting a call in the widest sense. This broadcast search component [33] is where the "open" part of the term finds foundation. The concept of the open call meets with a historical perception in crowdsourcing that the crowd participant's engagement with the process must be entirely of his or her own volition. This is the case in the vast majority of crowdsourcing portals where the crowd chooses a task. Strictly speaking employees of internal crowdsourcing initiatives can be said to participate on a non-voluntary basis where they are paid for their services. However, in such examples of internal crowdsourcing extra incentives are often presented to employees to secure their participation above ordinary remuneration. The following process components form the core of the crowd specification sub process and these hypotheses will be tested in subsequent stages of the research.

6.1 Crowd identification

Firstly, at the *crowd identification* stage the process initiator or crowd commences the process of identifying the attributes of the crowd membership required for the task. This can range from a broad open call for menial tasks to a very narrow requested skill set for specialist areas. In addressing characteristics of crowdsourcing processes Geiger identifies the pre selection of contributions as a first dimension of the process. To achieve this goal a crowd is selected. Different approaches are evident in practice as to how and when a crowd is selected as part of the initiative. With many examples an organization or individual has a clear idea in advance as to the skills required to complete a task. Where the skills required to complete the task are generic, tasks tend to be offered generally to a larger crowd without the need for substantial membership requirements. To the contrary where specific skills or qualifications are required to address a task the open call can be directed to a narrow or specific group of individuals with a very specific skill set. Software development and software testing is one such example of crowdsourcing where the initiator can select crowd membership based upon requisite skills [54, 15].

6.2 Crowd evaluation

Secondly, at the *crowd evaluation* stage the vetting of crowd members commences. In broad initiatives the users will subscribe to the site to commence participation. This is the case with most AMT tasks and projects on InnoCentive. In narrower initiatives

crowd users will post a profile and credentials in an effort to get selected for the task advertised. Curation is one such example of vetting. Prpic contextualizes this process in crowdsourcing whereby crowd members are "vetted in one way or another". The task of "curation" is often actuated through symbols or information in the information system directly associated with an individual's screen name and/or profile [44]. In essence the crowd can help select or include other crowd members from a community in a process based upon certain skillsets or profiles.

Furthermore, we can see cases where the crowd self selects other members at later points within ecosystems and tournaments. The portal SecondLife [37] presents many examples of crowds selecting other members for participation. It is posited that the process of "crowd selecting crowd" should not be confined to one stage in the process but can be completed multiple times over the course of a project depending upon the levels of task disaggregation completed and the concurrent skills required for the completion of tasks.

6.3 Crowd formulation

Lastly, at the *crowd formulation* stage the process initiator or crowd will formally confirm crowd membership for the applicant crowd user. This can either be through formal selection by an administrator or voting/selection through other crowd members. Furthermore, membership can be confirmed for a fixed period of time or number of process iterations. From practice we can see that the traditional crowd formulation process in many cases takes place at the start of the initiative. However, from literature we can see other examples of the crowd been formed in different capacities and at different stages of the process. Where examining social networks harnessed by the crowd one could argue that in many cases the crowd is not formed after the issuing of an open call but in many cases the community pre-exists the formation of the challenge. From this we deduce that the crowd can be formulated at any stage during the process. This concept runs contrary to the sequential nature of many concurrent crowdsourcing initiatives.

7. Cyclical iterative crowd processes

Dynamic examples exist of crowdsourcing initiatives whereby competition and collaboration are evident [1]. Such dynamic processes serve to assist the conceptual model advanced in this paper. The use of sprints as a process components are often a feature of agility and in particular within the area of agile software development. Methods such as scrum and R-scrum feature cyclical sprints as a core part of the development process [55, 12, 14]. Sprints are used with varying objectives and time frames in agile processes in an effort to maintain performance whilst continually refining the prototype or product under development. Cyclical work sprints are not exclusively confined to agile software development and can be witnessed within crowdsourcing. Little et al addresses the use of cyclical processes in crowdsourcing contests [36]. In one such experiment a cycle was applied to an image description task. A worker was provided with a human intelligence task (HIT) on AMT. They were shown a paragraph describing an image and were asked to improve the description. A second stage of the process displayed both original and the modified descriptions for the purposes of voting. A repeat cycle with a further advanced description is subsequently presented and voted upon [36]. Such iterative refinement cycles can be used in the process of crowd selection. A variety of approaches can be used whereby initial crowd members are selected who in turn go through similar cycles of crowd specification and selection.

Due to its menial nature a much broader crowd can be engaged where specialist knowledge or skills are required for active participation. Accordingly, the crowd specification process where cyclical in nature can through various cycles and employ the use of crowds of varying skill at various different stages depending on the complexity of the disaggregated task. Furthermore, different stages of the crowdsourcing process can be modularized into different cycles so as to further reduce process complexity. Baldwin states that modularization can allow modules to be worked on independently and in parallel, without intense on-going communication across modules [3]. How the process is broken down can have an effect on the type of crowd member required for a particular task of a certain size and level of complexity. However, for the purposes of the sub process advanced herein the level of task disaggregation / granularity selection is dealt with in another sub process in crowdsourcing.

A conceptual model is presented of the crowd specification sub-process that is not restricted from the outset and features a flexible means of refinement through iterative cycles. Secondly, crowd engagement is not static and typically operates in a changing environment, subjected to the same uncertain requirements and competitive pressures evident in contemporary product development processes. It is therefore argued that in following examples from other areas such as agile software development, the use of iterative sprints could be used to further refine crowd membership throughout a process. To date this is an issue not explicitly addressed in the existing literature.

8. Next stages of research

This research in progress forms part of a larger research initiative seeking to create an all encompassing crowdsourcing definition and framework. For the purposes of creating an overarching framework it is envisaged that this sub-process will be amalgamated into a full crowdsourcing framework in later research. Where completed an overarching crowdsourcing framework will facilitate the deployment in practice of a dynamic process possessing greater flexibility over existing linear and sequential crowdsourcing processes. The next steps of this research will involve the completion of three case studies in respect of the different task groupings identified from the literature. Thereafter, the model where validated can be replicated in testing of other parts of the crowdsourcing process. The completion of case studies as exemplars encompassing the three different stages of the crowd specification process is proposed for validation purposes. Having regard for the three separate processes identified within the crowd specification process a single case study was deemed insufficient by the authors. Multiple cases are suggested to increase the methodological rigor of the study through strengthening the precision, the validity and stability of the findings [40]. In selecting a multiple case approach guidelines as established by Yin were followed [63]. Case studies can be exploratory, descriptive, or serve explanatory purposes [51]. In the context of the present research it is proposed that the studies will primarily be explanatory in nature whereby the different parts of the sub process will seek to acquire validation in context. Rather than testing a theory or documenting a rare case the focus of the studies proposed concern analytical generalization where theory is to be reflected in the form of a sub process (model). Case study research has become increasingly popular as a method in IS and software engineering research [58, 4, 62, 41]. The case study methodology is particularly suited to studying real-world phenomena. Such instances cannot be studied separately from their context [48].

For this study we propose conducting face-to-face, semi-structured interviews with core personnel within crowdsourcing processes. It is proposed that interview data will be coded and placed under the themes identified in literature review process pertinent to the three stages of crowd specification. The first case study will deal with the validation of the *Crowd Identification* component of the process. To this end various studies are been considered including examples relating to the software development and software testing industry. In these examples crowd members with specific high quality skill sets are sought out in major portals. The second case study will focus upon the *Crowd Evaluation* stage and the mechanisms by which process initiators or crowd members access and evaluate the membership of the crowd. AMT has presented several examples of competition that are pertinent for such a type of case study. The final step of *Crowd Formulation* is of particular importance in innovation challenges. To this end several medical device providers have been identified for the purposes of examining the crowd formulation stage of the sub process.

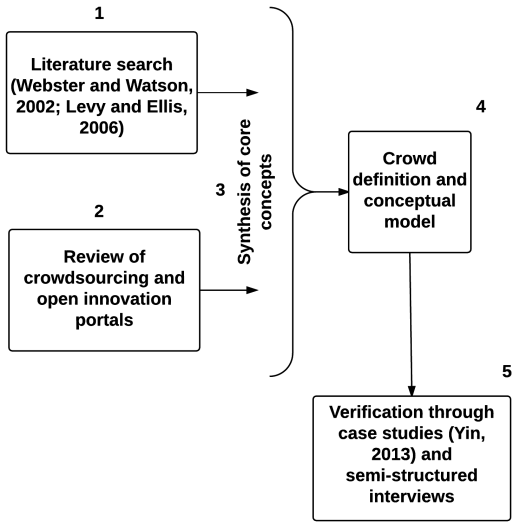

Figure 2. Proposed research method

9. Research limitations

The present research contains several limitations. Firstly, this paper addresses one aspect of the crowdsourcing process and not the process as a whole. Future research will seek to validate an overarching framework including the crowd specification components. Secondly, where it is envisaged that the conceptual model advanced herein can be deployed in existing crowdsourcing processes certain initiatives by nature would require substantial amendment to facilitate the inclusion of iterative processes. Further research is required to establish how the crowd specification process can be added to such traditional systems. Lastly, although identified through literature certain parts of the crowd specification process may not be used in every type of crowdsourcing initiative. Further research will be required to see if the model still functions where one part of the process is not utilized in the process of crowd specification. It is not possible from present research to hypothesize whether or not the model can function on an incomplete basis.

10. Conclusions

The conceptual model presented is grounded in two aspects of IS for the purpose of presenting theory. Firstly, the model is prescriptive in the sense that the model prescribes how crowd specification process should be done with the goal of attaining better results. The articulation of the means by which systems should be designed is an established concept in IS research [19]. Secondly, the model is instantiated and presented as a series of relationships amongst constructs that form part of one sub process within crowdsourcing projects. The identification of constructs and the relationship between them is an established form of theory presentation reflected in areas such as technology acceptance [9]. Such testable propositions will be examined at a later stage of the research through the use of case studies and semi-structured interviews. This research provides several contributions to the IS domain and the area of crowdsourcing. Firstly, a definition of the crowd has been advanced from constructs identified in literature. Secondly, a conceptual model is presented providing several advantages over traditional models in use. This model is a representation of theory derived from constructs identified from crowdsourcing research. Through the guise of this conceptual model the crowd specification process is no longer confined to the start of the process. Tailored crowds can be assembled and further refined at other stages of the process based upon changing skill requirements and needs of the tasks presented. This feature more closely aligns with the reality of crowdsourcing projects displaying varying and changing levels of task complexity and scale. Crowd initiatives are no longer reliant upon the use of a 'one size fits all' crowd across all tasks and challenges. The sub process presented in this paper facilitates the crafting of specific crowds for specific tasks through iterative processes. Under the model presented crowd participants can be utilized for selecting and evaluating the membership of other crowd participants. The crowd can be used to complete complex tasks such as evaluating prospective crowd member skills. Such a dynamic crowd enabled process can serve to reduce costs for businesses and organizations engaging the crowd. Future stages of the research will seek to address similar challenges faced in other parts of the crowdsourcing process.

Crowd specification forms one important part of the overall process. However, other tasks and challenges such as defining the problem to be addressed through the crowdsourcing project form sub processes in their own right. The format of this research can be applied in tandem to other parts of the process. A similar research approach could form the means through which the remainder of an overarching framework could be addressed. From the research herein we conclude that crowdsourcing processes can feature a dynamic crowd who are capable of completing complex tasks such as self-selection and crowd vetting. The full extent of this dynamism will be tested in future stages of this research.

11. Acknowledgements.

This work was supported, in part, by Science Foundation Ireland grant 10/CE/I1855 and by the Irish Research Government of Ireland 2013 Research Project Grants Scheme Council - RPG2013-6 (SFI/HEA).

References

[1] A. Afuah and C. L. Tucci, "Crowdsourcing As a Solution to Distant Search", Academy of Management Review, 37 (2012), pp. 355-375.

[2] D. Anastasiou and R. Gupta, "Comparison of crowdsourcing translation with Machine Translation", Journal of Information Science, 37 (2011), pp. 637-659.

[3] C. Baldwin and E. V. Hippel, *Modeling a Paradigm Shift : From Producer Innovation to User and Open Collaborative Innovation Modeling a Paradigm Shift : From Producer Innovation to User and Open Collaborative Innovation 1 . Introduction and Overview*, 2010, pp. 1-37.

[4] I. Benbasat, D. K. Goldstein and M. Mead, "The case research strategy in studies of information systems", MIS quarterly (1987), pp. 369-386.

[5] I. Benbasat and R. Zmud, "Empirical Research in Information Systems: The Practice of Relevance", MIS Quarterly, 23 (1999), pp. 3-16.

[6] D. C. Brabham, *Using Crowdsourcing In Government*, *IBM Center for The Business of Government*, 2013.

[7] H. W. Chesbrough, "The Era of Open Innovation", Sloan Management Review, 44 (2003), pp. 35-41.

[8] J. R. Corney, C. T. Sanchez, A. P. Jagadeesan and W. C. Regli, "Outsourcing labour to the cloud", International Journal of Innovation and Sustainable Development, 4 (2009), pp. 294-294.

[9] F. D. Davis, R. P. Bagozzi and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models", Management science, 35 (1989), pp. 982-1003.

[10] D. DiPalantino and M. Vojnovic, *Crowdsourcing and all-pay auctions*, ACM Press, New York, New York, USA, 2009, pp. 119-119.

[11] A. Doan, R. Ramakrishnan and A. Y. Halevy, "Crowdsourcing systems on the World-Wide Web", Communications of the ACM, 54 (2011), pp. 86-86.

[12] B. Düring, *Sprint Driven Development: Agile Methodologies in a Distributed Open Source Project (PyPy)*, *Extreme Programming and Agile Processes in Software Engineering*, Springer, 2006, pp. 191-195.

[13] E. Estelles-Arolas and F. Gonzalez-Ladron-de-Guevara, "Towards an integrated crowdsourcing definition", Journal of Information Science, 38 (2012), pp. 189-200.

[14] B. Fitzgerald, K.-J. Stol, R. O'Sullivan and D. O'Brien, *Scaling agile methods to regulated environments: An industry case study*, *Proceedings of the 2013 International Conference on Software Engineering*, IEEE Press, 2013, pp. 863-872.

[15] D. Fried, "Crowdsourcing in the software development industry", Nexus of Entrepreneurship and Technology Initiative Fall (2010).

[16] R. D. Galliers, *Relevance and rigour in Information Systems Research: Some Personal Reflections on Issues Facing the Informations Systems Research Community*, in B. C. Glasson, I. T. Hawryszkiewycz, B. A. Underwood and R. A. Weber, eds., *Business Process Re-Engineering: Information Systems Opportunities and Challenges*, 1994, pp. 93-101.

[17] D. Geiger, S. Seedorf, T. Schulze, R. C. Nickerson and M. Schader, *Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes*, *Americas Conference on Information Systems (AMCIS)*, Detroit, Michigan, 2011.

[18] D. Govindaraj, K. V. M. Naidu, A. Nandi and G. Narlikar, "MoneyBee : Towards Enabling a Ubiquitous , Efficient , and Easy-to-Use Mobile Crowdsourcing Service in the Emerging Market", Bell Labs Technical Journal, 15 (2011), pp. 79-92.

[19] S. Gregor, "The Nature of Theory in Information Systems", MIS Quarterly, 30 (2006), pp. 611-642.

[20] A. Gupta, W. Thies, E. Cutrell and R. Balakrishnan, *mClerk : Enabling Mobile Crowdsourcing in Developing Regions*, 2012, pp. 1843-1852.

[21] A. Halfaker and J. Riedl, "Bots and cyborgs: Wikipedia's immune system", Computer (2012), pp. 79-82.

[22] J. Hamari, J. Koivisto and H. Sarsa, *Does gamification work?--a literature review of empirical studies on gamification*, *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, IEEE, 2014, pp. 3025-3034.

[23] L. Harris, *How you can develop a Skunk Works for your organization*, I E E E, New York, 1997.

[24] D. Hirshleifer and S. Hong Teoh, "Herd behaviour and cascading in capital markets: A review and synthesis", European Financial Management, 9 (2003), pp. 25-66.

[25] M. Hossain, *Users' motivation to participate in online crowdsourcing platforms*, *Innovation Management and Technology Research (ICIMTR), 2012 International Conference on*, IEEE, 2012, pp. 310-315.

[26] B. J. Howe, *The Rise of Crowdsourcing*, *Wired*, 2006, pp. 1-5.

[27] IBM, *IBM Jam Events*, 2015.

[28] C. E. Idicheria, A. Schellong, J. Fiedler, M. Kammer, M.-T. Huppertz and H. Westerfeld, *A Review of 311 in New York City*, Technical report, 2012.

[29] L. B. Jeppesen and K. R. Lakhani, "Marginality and problem-solving effectiveness in broadcast search", Organization science, 21 (2010), pp. 1016-1033.

[30] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey and S. Linkman, "Systematic literature reviews in software engineering–a systematic literature review", Information and software technology, 51 (2009), pp. 7-15.

[31] A. Kittur, E. Chi and B. Suh, "Crowdsourcing for usability: Using micro-task markets for rapid, remote, and low-cost user measurements", Proc. CHI 2008 (2008).

[32] G. Klein, J. J. Jiang and C. Saunders, "Leading the Horse to Water", Communications of the AIS, 18 (2006), pp. 259-274.

[33] K. R. Lakhani, *Broadcast Search in Problem Solving: Attracting Solutions from the Periphery1*, *Technology Management for the Global Future, 2006. PICMET 2006*, IEEE, 2006, pp. 2450-2468.

[34] E. Law and L. v. Ahn, "Human Computation", Synthesis Lectures on Artificial Intelligence and Machine Learning, 5 (2011), pp. 1-121.

[35] Y. Levy and T. J. Ellis, "A systems approach to conduct an effective literature review in support of information systems research", Informing Science: International Journal of an Emerging Transdiscipline, 9 (2006), pp. 181-212.

[36] G. Little, L. B. Chilton, M. Goldman and R. C. Miller, *Turkit: tools for iterative tasks on mechanical turk*,

Proceedings of the ACM SIGKDD workshop on human computation, ACM, 2009, pp. 29-30.

[37] G. Llewelyn, *Crowdsourcing in Second Life*, in G. Llewelyn, ed., *Gwyneth Llewelyn's essays on Second Life® socio-economic issues.*, 2006.

[38] A. Majchrzak and A. Malhotra, "Towards an information systems perspective and research agenda on crowdsourcing for innovation", The Journal of Strategic Information Systems, 22 (2013), pp. 257-268.

[39] C. McPhail and R. T. Wohlstein, "Individual and collective behaviors within gatherings, demonstrations, and riots", Annual Review of Sociology (1983), pp. 579-600.

[40] M. B. Miles and A. M. Huberman, *Qualitative data analysis: An expanded sourcebook*, Sage, 1994.

[41] A. Mockus, R. T. Fielding and J. D. Herbsleb, "Two case studies of open source software development: Apache and Mozilla", ACM Transactions on Software Engineering and Methodology (TOSEM), 11 (2002), pp. 309-346.

[42] C. Okoli and K. Schabram, "A guide to conducting a systematic literature review of information systems research", (2010).

[43] M. K. Poetz and M. Schreier, "The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?", Journal of Product Innovation Management, 29 (2012), pp. 245-256.

[44] J. Prpic and P. Shukla, *The Contours of Crowd Capability*, *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, 2014, pp. 3461-3470.

[45] J. Prpić, P. P. Shukla, J. H. Kietzmann and I. P. McCarthy, "How to work a crowd: Developing crowd capital through crowdsourcing", Business Horizons, 58 (2015), pp. 77-85.

[46] A. J. Quinn and B. B. Bederson, *Human computation: a survey and taxonomy of a growing field*, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2011, pp. 1403-1412.

[47] M. Rosemann, "Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks", MIS Quarterly, 32 (2008), pp. 1-22.

[48] P. Runeson, M. Host, A. Rainer and B. Regnell, *Case study research in software engineering: Guidelines and examples*, John Wiley & Sons, 2012.

[49] G. D. Saxton, O. Oh and R. Kishore, "Rules of Crowdsourcing: Models, Issues, and Systems of Control", Information Systems Management, 30 (2013), pp. 2-20.

[50] E. Schenk and C. Guittard, "Towards a characterization of crowdsourcing practices", Journal of Innovation Economics & Management (2011), pp. 93-107.

[51] M. Shakir, "The selection of case studies: strategies and their applications to IS implementation case studies", (2002).

[52] H. Simula and M. Vuori, "Benefits and barriers of crowdsourcing in B2B firms: Generating ideas with internal and external crowds", International Journal of Innovation Management, 16 (2012), pp. 1240011.

[53] N. J. Smelser, *Theory of collective behavior*, Quid Pro Books, 2011.

[54] K.-J. Stol and B. Fitzgerald, *Two's Company , Three's a Crowd : A Case Study of Crowdsourcing Software Development*, 2014.

[55] K. Vlaanderen, S. Jansen, S. Brinkkemper and E. Jaspers, "The agile requirements refinery: Applying SCRUM principles to software product management", Information and software technology, 53 (2011), pp. 58-70.

[56] M. Vukovic, *Crowdsourcing for Enterprises*, *Services-I, 2009 World Conference on*, IEEE, 2009, pp. 686-692.

[57] B. Walsh, C. Maiers, G. Nally and J. Boggs, "Crowdsourcing individual interpretations: Between microtasking and macrotasking", Literary and Linguistic Computing, 29 (2014), pp. 379-386.

[58] G. Walsham, "Interpretive case studies in IS research: nature and method", European Journal of information systems, 4 (1995), pp. 74-81.

[59] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review", Management Information Systems Quarterly, 26 (2002), pp. 3.

[60] J. Yang, L. A. Adamic and M. S. Ackerman, *Crowdsourcing and knowledge sharing; strategic user behavior on taskcn*, ACM Press, New York, New York, USA, 2008, pp. 246-246.

[61] J. Yap, *Asian firms well-poised for internal crowdsourcing*, *ZD Net*, 2012.

[62] R. K. Yin, "Applications of case study research (applied social research Methods)", Series, 4th edn. Thousand Oaks: Sage Publications (2003).

[63] R. K. Yin, *Case study research: Design and methods*, Sage publications, 2013.