A Conceptual Model to Understand the Factors that Drive Individual Participation in Crowdsourcing for Medical Diagnosis

Kaushik Ghosh
Lamar University
kghosh@lamar.edu

Kabir Sen
Lamar University
kabir.sen@lamar.edu

Abstract

In healthcare, assisted by collective knowledge of a large group of individuals, crowdsourcing is enabling prediction of disease outbreaks and diagnosis of rare medical conditions. This article is a 'Research-in Progress' that examines the role of Web-based platforms in driving participation of individuals in crowdsourcing services for medical diagnosis. Based on existing literature, a conceptual research model is developed that outlines factors that drive individual participation. A set of propositions based on the research model are developed. A research methodology is proposed with a plan for the empirical analysis. Contributions and implications are discussed.

1. Introduction

Crowdsourcing is defined as the outsourcing of a task to a “crowd”, or a population of unknown individuals, rather than to an identified person, group or team, in form of an open call. Compared to the more traditional practice of depending on individual expertise, crowdsourcing uses the collective wisdom of individuals to solve problems [11, 24]. An individual (hereafter referred to as contributor) from the crowd self-selects to find a solution to the problem, often in exchange for monetary rewards offered by the individual (hereafter referred to as seeker) that submits the crowdsourcing task request. Although crowdsourcing has been around since the eighteenth century, the ubiquity and ‘power’ of the Internet have enabled seekers to virtually broadcast problems through Web-based information systems to unknown individuals (contributors) located anywhere in the world. Potential contributors, irrespective of her geographical location, can self-select to participate in crowdsourced challenges [1].

Businesses have used crowdsourcing in diverse domains to achieve a variety of goals [24]. For example, Unilever leverages the intellect of the crowd to seek ideas to reduce sodium content in food [2]. GE reaches out to the crowd to get ideas on how to make more energy efficient devices [17]. In healthcare, crowdsourcing is employed to accomplish a variety of tasks; medical transcription, diagnosis and treatment of rare diseases, perform medical research, and predict outbreaks of infectious diseases [22], among others. Many healthcare experts have touted crowdsourcing to be a model for the future to meet the growing challenges of rising costs and lack of quality healthcare [10].

Popular press calls crowdsourcing a technology-based innovation that could change the landscape of medical diagnosis [16]. Indeed, there are several examples of popular crowdsourcing initiatives for disease treatment and diagnosis. A research group at UCLA has developed a game that attracts crowds who help to identify cells infected with malaria, a disease that affects over half a billion people worldwide, every year [18]. The New England Journal of Medicine has a routine feature requesting readers to help diagnose unexplained medical conditions based on photographs posted by editors of the journal [14]. CrowdMed, a recently launched Web-based crowdsourcing platform uses the wisdom of the crowd to diagnose rare medical conditions that doctors are unable to uncover [18].

For this paper, we focus on crowdsourcing in context of healthcare and provide a narrow definition of crowdsourcing that involves tasks of diagnosing rare medical conditions by contributors. The seekers virtually broadcast narratives of their health issues via a Web-based platform to obtain a medical consultation. While a contributor is an individual who self-selects to contribute to the crowdsourcing task, a seeker is the entity seeking to obtain a diagnosis or cure for her disease condition.

It may be noted that the crowdsourcing platform acts as an agent or intermediary overseeing the exchange between the seeker and the contributor, and managing the crowdsourcing activity or process. Further, the current study limits its scope to crowdsourcing services, similar to Crowdmed, which utilize the collective input of individuals, select the most promising solutions for treating the health condition, and broadcast it to the particular patients seeking medical consultations.

Stimulating contributor participation in crowdsourcing tasks poses challenges [6]. Studies [for e.g., 3, 13] have suggested that a set of intrinsic and
extrinsic motivations grounded in classical motivation theory may drive potential contributors to participate in crowdsourcing. Further, in crowdsourcing, the seeker usually is an enterprise that elects to outsource a problem rather than solve it internally or designate a contractor for the task, after assessing that crowdsourcing is the best alternative [1]. However, the seeker’s decision to select crowdsourcing as an option may depend on the contextual nature of problem [1] – which in the case of this study, is finding cure for a rare medical condition. Therefore, a seeker’s decision to share their health problem with unknown individuals to find cure for their condition may depend on set of factors that differ from that of an organization seeking solution to a business problem (e.g. marketing idea for a product), and needs to be examined. Moreover, past research [27] has suggested that the crowdsourcing platform characteristics may play a significant role in driving the motivations of contributors and seekers. A platform refers to the Web-based setting that is used to transfer, receive, store, or process information relevant to the crowdsourcing process [1]. As stated earlier, crowdsourcing platform acts as the intermediary and manages the crowdsourcing process - ensuring that requests are communicated to seekers, issuing authentication credentials to them, overseeing any misconduct, monitoring the credibility of the solutions, and rewarding the contributors [33].

This study proposes a research model investigating the following: (1) what attracts seekers to crowdsourcing initiatives to obtain consultation for their health problems; (2) what attracts contributors to participate in crowdsourcing task that involves diagnosing a rare medical condition or disease; (3) how the Web-based platform control system influences seeker and contributor motivations to participate in crowdsourcing.

The following section summarizes prior research on crowdsourcing in healthcare, outlines the motivations of seekers and contributors, and discusses the role of Web-based crowdsourcing platform. Then, we present a research model for participation in the context of crowdsourcing for medical diagnosis and develop propositions. Next, we discuss methods for data collection and suggest techniques to test the proposed model. The paper concludes with implications for both research and practice.

### 2. Prior Research

#### 2.1. Crowdsourcing in Healthcare

Crowdsourcing is often considered as a virtual production and collaboration platform for enabling collective idea generation and problem solving [25]. In healthcare, crowdsourcing powered by the Internet and assisted by collective knowledge of large-groups of individuals, is supporting health-related research collaboration, enabling prediction of disease outbreaks, and allowing diagnosis of rare medical conditions [30]. The medical vocation is known for its long and demanding years of study, an increasing trend towards specialization and strict regulations. However, the challenges currently faced by healthcare providers, patients and medical researchers seem ideal to the collective problem-solving model afforded by crowdsourcing. There are several examples of crowdsourcing initiatives currently underway in the healthcare domain, some of which have been included in Table 1.

<table>
<thead>
<tr>
<th>Crowdsourcing Platform/Domain</th>
<th>Activity &amp; Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>CrowdMed/Medical Diagnosis [30]</td>
<td>Patients with rare diseases list their case to seek consultation for possible diagnosis from contributors who are not necessarily medical experts. Contributors provide ideas for curing rare diseases.</td>
</tr>
<tr>
<td>Groundhog/Medical Diagnosis [30]</td>
<td>Physicians post medical cases to seek help from peers. Doctors are directed to appropriate diagnosis, reducing the number of unnecessary tests, ineffective treatments, and consultations.</td>
</tr>
<tr>
<td>Webicina/Medical Research [30]</td>
<td>Doctors worldwide use social media to share and communicate their research findings. Collaboration and communication of new findings in medical research is disseminated rapidly.</td>
</tr>
<tr>
<td>PatientsLikeMe/Medical Diagnosis [32]</td>
<td>Patients with certain rare health conditions share and compare their symptoms and responses to different treatments with individuals who have had similar experiences.</td>
</tr>
</tbody>
</table>

Table 1. Examples of Crowdsourcing in Healthcare

A major advantage of crowdsourcing in healthcare is that it allows aggregation of information and ideas contributed by large groups of individuals from diverse backgrounds and experiences. This collective wisdom may possibly provide solutions to problems that have confounded doctors and scientists. The promise of
services such as CrowdMed is that, it can effectively organize the collection and distribution of medical wisdom from diverse source of contributors and aim to distribute the more promising ones to its group of seekers. Healthcare experts believe that leveraging the knowledge of crowd can reduce rising healthcare costs, make patient care more efficient, and save lives [33].

2.2. Seekers Motivations for Participation in Crowdsourcing

Seekers can search for health-related information from various sources, including face-to-face channels (doctor/care provider) and online channels such as the Internet [36]. Common methods to seek information online include looking up information using a search engine, emailing a query to a specific person, or broadcasting a question to a social networking site (such as Facebook, Twitter) [8, 20]. For this current study, we define health-information seeking as the act of crowdsourcing a health-related issue, or looking for health (or medical) consultation, by broadcasting health-related question(s) through a Web-based platform to unknown individuals to obtain cure for a disease condition.

Search for information is rooted in the concept of information-seeking behavior theory [8], which exemplifies need for information to achieve a certain goal as the primary motivation that drives the search process. In context of this current study, an individual (the seeker) may seek online consultation from a crowd of people to obtain a diagnosis or find a cure for her disease. A seeker’s need to look for online health information or crowdsource a health-related issue may be related to factors, such as value for personal health and health locus of control [15, 36]. Health value of an individual refers to the importance she places on her health [26]. An individual who values health highly will not only seek more information, but will also explore all available health information channels to improve her condition [35]. Health locus of control is the belief in one’s own ability to take control of her health condition [36]. An individual with a higher health locus of control tends to search more for health information [5].

Further, information seeking may also depend on an individual’s (or seeker’s) situational factors related to the disease, such as severity of the disease condition [36], defined as the extent to which the disease poses a serious threat to the individual’s life. For example, certain types of cancers (such as stage IV Leukemia) are life threatening, while diabetes, a chronic (long-lasting) condition does not present a threat to one’s life. In our conceptualization of disease severity, Leukemia is considered a severe health condition, while, diabetes is not. Another disease-related factor that may influence information seeking behavior is the disease treatment knowhow, which refers to the information available to the patient regarding her disease treatment, through traditional sources (doctors/care providers). A patient will be more inclined to seek crowdsourcing avenues to obtain information related to cure for her disease if the knowhow related to the disease is little or absent.

2.3. Contributors Motivations for Participation in Crowdsourcing

Regarding contributors’ motivations to self-select in a crowdsourcing activity, prior research has studied the topic in the context of open source projects, a form of crowdsourcing to build software programs [28], contests involving creation of Website designs [37], online innovation communities [21], tasks that involve medical transcription [13], and activities that include generation of ideas for a new product [23]. These studies applied motivation theory and suggested that a variety of intrinsic and extrinsic factors determine contributor participation in crowdsourcing. Commonly cited intrinsic factors included a sense of helping a community [23], satisfaction of solving a challenging problem [21], and personal gratification [28]. Frequently mentioned extrinsic factors were earning monetary rewards and name recognition [37].

It has been suggested that many of these motivating factors are dependent on the nature of the problem to be solved or type of activity undertaken [1]. In other words, contributors in a crowdsourcing activity involving idea generation for a new product may have a different set of motivations, when compared to that of participants that are involved in crowdsourcing to diagnose a disease. Contributors self-selecting in the new product development task may be driven by satisfaction of being involved in innovation [1]; whereas those individuals providing inputs to help disease diagnosis may be motivated by a sense of obligation to patients suffering from the disease. Further, we contend that certain motivating factors that matter in crowdsourcing for new product innovation could be of consequence in crowdsourcing involving disease diagnosis; in both groups, contributors’ reason to self-select could be influenced by their expectation of monetary rewards.

In context of this study - crowdsourcing tasks that involve diagnosis of rare disease conditions, we draw on prior research and arguments specified in motivation theory and suggest a set of intrinsic and extrinsic factors that may have a significant influence on contributor’s choice to participate. We posit that intrinsic factors, namely, community affection defined
as the extent to which an individual identifies with the purpose and goal of the specific crowdsourcing initiative [23], and personal gratification, defined as the feeling or sense of satisfaction of contributing to crowdsourcing task [13]; as well as, extrinsic factors, namely, expectations of monetary rewards and name recognition [37], may determine contributor’s participation in crowdsourcing.

2.4. Web-based Platform Control Systems

The concept of crowdsourcing is not new, and has existed for over hundred years. The Internet and communication technologies of the twenty-first century have proliferated crowdsourcing to the extent that it has become a tool to find solutions to problems that may not otherwise be resolved using traditional approaches [29, 33].

Web-based platforms refer to information systems that assimilate human and computational agents to facilitate the process of outsourcing a task and aggregating ideas from the crowd [9]. Challenges of crowdsourcing include attracting seekers and contributors to the platform, having an evaluation system to monitor the quality and credibility of the contributions (ideas), assuring compensation or reward to the contributors,) and maintaining a “balance” between “openness” and privacy [6]. Further, specific to crowdsourcing tasks involving medical diagnosis, storage, presentation, and retrieval of information (for example, patient medical history, digital images or videos illustrating a medical condition, such a mysterious skin abrasion) essential to solve the crowdsourcing challenge, must be supported by features built into the platform.

In the Internet-enabled crowdsourcing model, the Web-based platform performs several salient functions [24, 26]. It serves as the online environment were seekers can upload and post details about the task; advertises the task on behalf of seekers; authenticates the identity of seekers and contributors; handles execution of the crowdsourcing requests; provides technical tools to contributors to undertake the task; establishes the rules to be followed by participants; analyzes the possible solutions offered by the contributors and validates them to provide the optimal remedy to the seeker; and compensates the deserving contributor [33]. In principle, functions embedded in the Web-based platform provide strategic value-addition to the crowdsourcing model. Further, for crowdsourcing platform to support these salient functions, control systems have to be in place. Control systems are defined as the policies, procedures and mechanisms built into the crowdsourcing platform that integrate individual and computational agents to enable the process of outsourcing and collecting contributions from a large group of individuals, and presenting optimal solutions to seekers [25]

3. Conceptual Framework

A conceptual research framework for participation of seekers and contributors in crowdsourcing for medical diagnosis is presented in Figure 1. This framework integrates research issues at the intersection of information seeking, work motivation, and Web-based system characteristics. The focus of this research is how the interaction of the seeker and the contributor with the crowdsourcing platform determines participation in crowdsourcing. In the context of crowdsourcing for medical diagnosis, suggestions for cure and diagnosis offered may complement the consultations from doctors/physicians. Underlying the conceptual framework, the major motivation variables are grouped under personal factors, disease-related factors, intrinsic factors, and expectations. The set of platform control system characteristics that enable participation are described as the policies, procedures, and mechanisms built into the Web-based environment of the crowdsourcing platform. The framework helps explore the moderating role of platform control systems in enhancing participation of seekers and contributors in the crowdsourcing task.

### Figure 1: Model of Participation in Crowdsourcing for Medical Diagnosis

#### 3.1. Propositions

Health information needs embrace both cognitive needs, which include information for disease prevention and treatment, and affective needs [19], which include information for coping with disease.
Based on the information-seeking behavior paradigm, a seeker is likely to use crowdsourcing service when she has high health value and high health locus of control [37]. Indeed, personal health orientation could predict a person’s preference toward health information channels. The individual who has high ‘internal’ values related to health will seek more information than one who does not value health or who holds external beliefs.

Further, information seeking is also motivated by information needs, a perception tied to knowledge insufficiency. In addition, when the individual is faced with a disease that is life threatening and traditional methods for treatment have not been successful, individual will resort to seeking help from crowdsourcing community [37]. Thus, the following two propositions have been developed.

**Proposition 1a:** Personal factors (health value and health locus of control) related to the seeker will determine her intention to participate (use) in the crowdsourcing service.

**Proposition 1b:** Disease-related factors (disease severity and disease knowledge) of the seeker will determine her intention to participate (use) in the crowdsourcing service.

Anchoring to the views embraced in motivation theory, a contributor’s inclination to participate in crowdsourcing is based on her affection for community and personal gratification (intrinsic factors), and the monetary rewards and recognition (extrinsic factors), she expects from participation in crowdsourcing.

Community affection, also referred to as community identification relates to the idea of an individual’s affinity to a group. It has been linked to an individual’s intention to participate in virtual communities, such as blogs [20], and open source projects [28]. Another aspect of intrinsic motivation – personal gratification pertains to the work engagement by an individual, because the work itself is interesting or satisfying to her [37].

Expectations drive individual participation in OSS development [28]. Crowdsourcing studies [for e.g. 1, 4] have found different types of expectations to explain an individual contributor’s participation in a crowdsourcing contest. In this study, we focus on two types of expectations — motivation to gain monetary reward and motivation to gain recognition [37]. To attract the best solution for disease-related problems, the sponsor of a crowdsourcing activity may provide incentives to contributors in form of monetary rewards. Contributors’ intend to compete for such rewards may drive their participation in the crowdsourcing task.

Further, for contributors, participating in a crowdsourcing contest, it is an opportunity to draw the attention of seekers, as well as sponsors, and gain peer recognition for their expertise. Thus, the following two propositions have been developed.

**Proposition 2a:** Factors that are intrinsic to the contributor (community affection and personal gratification) will determine her participation in crowdsourcing for tasks involving disease diagnosis.

**Proposition 2b:** Contributor’s expectations of rewards and recognition (from the crowdsourcing activity) will determine her participation in crowdsourcing for tasks involving disease diagnosis.

In the context of our research model, technology built into crowdsourcing control system refers to the capabilities (policies, procedures, and mechanisms) or key functions that enable the crowd to form, disseminate pertinent information relevant to the crowdsourcing activity. It also includes authenticating participant identity, enabling interactions between the seekers and the contributors, facilitating trust building between them, analyzing and validating plausible solutions to the crowdsourcing problem, and driving the ‘ultimate’ solution [25, 31].

Based on prior research [e.g., 27], we argue that the crowdsourcing platform’s control system policies, procedures, and mechanisms will influence the motivations of the seekers and contributors to participate in the crowdsourcing service.

Crowdsourcing platforms that are perceived to have efficient mechanisms to validate solutions and provide ‘optimal’ suggestions to a specific patient (seeker) will positively moderate the relationship between the seeker’s personal factors and her intention to participate. Similarly, efficient management of compensation will positively moderate the relationship between contributor’s expectations of rewards and her participation. Thus, we contend that.

**Proposition 3a:** Platform control system will moderate the relationship between personal factors related to the seeker and her participation in crowdsourcing.

**Proposition 3b:** Platform control system will moderate the relationship between disease-related factors of the seeker and her participation in crowdsourcing.

**Proposition 4a:** Platform control system will moderate the relationship between intrinsic factors of
the contributor and her participation in crowdsourcing for tasks involving disease diagnosis.

Proposition 4b: Platform control system will moderate the relationship between contributor’s expectations (from the crowdsourcing activity) and her participation in crowdsourcing for tasks involving disease diagnosis.

4. Proposed Methodology

We propose to use a survey to collect data through a questionnaire survey of individuals involved in crowdsourcing. The questionnaire will be developed through reviewing existing literature. With respect to items for the construct, Platform Control System, because there are no established scales, we will develop new items by following the process prescribed in prior studies. An initial instrument for testing the research model (excluding items for Platform Control System) is included in the Appendix. Conceptualizations and definitions of the variables in our research model have been noted earlier in the paper.

The initial survey instrument items will be reviewed by a focus group that has participated in crowdsourcing activity. Items will be reworded to reflect their inputs. Because we need to develop new items for Platform Control System, we will incorporate inputs from the focus group interviews to formulate items for this construct.

A pilot test will be conducted using a small sample (n=75) from the target population consisting of both seekers and contributors, as a basis for further refinement of the survey instrument. Items will be revised or dropped, based on the results of the pilot study. This process will be continued until no further modifications to the questionnaire are necessary. Content validity will be verified by assessing content validity ratios for each scale item.

The final instrument will be deployed as an online survey which will be used to collect data from a large sample (n=300) participating in crowdsourcing services such as CrowdMed. Similar strategies for data collection have been used in prior studies on crowdsourcing [e.g. 13]. The empirical analysis will use structural equation modeling (SEM) techniques and/or other suitable estimation methods. We will include control variables (such as age, gender, education level etc.) that might influence intention to participate in crowdsourcing, beyond those proposed in our model. Currently, we are in the process conducting focus group interviews, following which we will be pilot testing the survey. We hope to present initial empirical evidence at the conference.

5. Discussion

The conceptual model suggested in this paper details a set of variables to explore what drives seeker and contributor participation in crowdsourcing involving medical diagnosis. Further, the model suggests that functions built into the crowdsourcing platform influences success of the crowdsourcing service, which is the focus guiding this study.

Two key theoretical contributions can be drawn from this study. First, the proposed model suggests that Web-based platforms functions may have a significant role to play in medical diagnosis crowdsourcing. Second, we contribute to existing literature on the emerging trends in Web-based tools for healthcare [7], by arguing that Web-based crowdsourcing model has the potential to make radical transformations to the traditional methods of medical treatment, by leveraging the collective wisdom of a large group of individuals. In terms of limitations of this research, the proposed framework needs to be empirically tested which is a future scope of this study.

6. Conclusion

It has been argued that compared to individual expertise, the promise of crowdsourcing services lies in the potential to collect wisdom from many sources [12], choose the most promising remedies, and disseminate it to specific patients. However, crowdsourcing services in healthcare may have some limitations. For example, lack of access to longitudinal medical data or disease history specific to a patient (seeker) may limit the ability of contributors to diagnose the disease correctly. Thus, in addition to finding ways to attract individuals to participate in crowdsourcing, special emphasis must be given while incorporating design ‘elements’ in the crowdsourcing platform so that it has the capacity to collect all possible historical medical data from the patient (seeker) and broadcast it to contributors. Further, the success of crowdsourcing in medicine should be judged not by comparing it to results obtained from consulting medical experts, but on how it has solved medical problems, which were once considered untreatable or incurable by most experts in the field of medicine.
7. References


8. Appendix

**Health Value** (refers to the importance an individual places on her health)

With regards to my own health:

1. *If my health condition is not good, I do not have anything.*
2. *There is nothing more I care about.*
3. *It is of major importance in my life.*
4. *There are a few things more important.*

**Community Identification** (perception about belonging to a community)

I participate in the crowdsourcing activity because:

1. *I am proud to belong to the crowdsourcing initiative.*
2. *I feel a sense of belonging to the crowdsourcing initiative.*
3. *I care about the long-term success of the crowdsourcing initiative.*
4. *I am loyal patron of the crowdsourcing initiative.*

**Disease Severity** (the extent to which the disease poses a serious threat to the individual’s life)

My disease:

1. *Is a serious condition*
2. *Has major consequences on my life*
3. *Does not have much effect on my life* (reverse coded)
4. *Has serious financial consequences*

Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree

**Disease Know-how** (the information available to the patient regarding her disease treatment, through traditional sources (doctors/care providers))

I can get:

1. *Access to information related to my disease anytime.*
2. *Access to information related to my disease wherever I go.*
3. *Access to information related to my disease just in time.*

Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree

**Health Locus of Control** (belief in one’s own ability to take control of her health condition)

As far as I am concerned:

1. *If I get sick, it is my own behavior, which determines how soon I get well again.*
2. *I am in control of my health.*
3. *I am directly responsible for my health.*
4. *The main thing that affects my health is what I myself do.*
5. *If I take care of myself, I can avoid illness.*
6. *If I take the right actions, I can stay healthy.*

Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree

(Adopted from: [35])

### Disease Severity

**Health as a Value: Theoretical and Methodological Considerations**, Health Psychology 5, 1, 1986, pp. 25-43)

**Personal Gratification** (sense of satisfaction of contributing to crowdsourcing task)

I participate in the crowdsourcing activity because:
1. It gives me the satisfaction of seeing the results.
2. It gives me the chance to do things I am good at.
3. I really enjoy it.
4. It gives me a sense of personal achievement.
Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree


**Monetary Rewards**

I participate in the crowdsourcing activity because:
1. I am paid for my contributions.
2. I receive some form of explicit compensation (e.g., salary).
3. For me, it is extremely profitable/not profitable.
Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree


**Recognition**

I participate in the crowdsourcing activity because:
1. I am recognized for my efforts.
2. My expertise is acknowledged
3. It gives me a sense of personal achievement.
Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree


**Intention to Participate**

1. I intend to participate in the crowdsourcing activity
2. I will try to participate in the crowdsourcing activity.
3. I am determined to participate in the crowdsourcing activity.
Suggested scale: 5-point Likert scale where 1=strongly disagree and 5= strongly agree

*(Adopted from: [37])*