American Journal of Information Systems, 2018, Vol. 6, No. 1, 13-22
Available online at http://pubs.sciepub.com/ajis/6/1/3

DOI:10.12691/ajis-6-1-3

An Infrastructural Analysis of a Crowdsourcing Tool for Environmental Research

Selamawit Molla Fossum[1,*], Susana Lopez-Aparicio[2], Håvard Vika Røen[2]

[1]Department of Informatics, University of Oslo, Oslo, Norway
[2]Norwegian Institute for Air Research Institute, kjeller, Norway
*Corresponding author: selama.molla@gmail.com

Abstract In this paper, we adopt information infrastructure design principles and concepts from the theory of critical mass to analyze and evaluate the socio-technical conditions that hindered the successful bootstrapping processes of a crowdsourcing tool for environmental research. The crowdsourcing tool was designed to improve the estimation of emissions from burning wood for residential heating in urban areas in Norway by collecting geolocation data on wood consumption and stove types. Our analysis identifies three groups of users, namely scientists, wood consumers (end users), and key stakeholders, that the IT capability of the tool needs to support. At this stage, we determined that the tool was more useful to the scientists than the other two groups, which was attributed to its low uptake. We uncovered various underlying issues through further analysis of means by which the tool becomes useful to key stakeholders. One particular issue concerned the tension between existing data collection practices, which are based on statistical methods, and the nature of crowdsourcing, which is based on the principle of open call with no sampling techniques. From our analysis, we concluded that developing crowdsourcing tools for research requires increasing the tool's benefits for key stakeholders by addressing these underlying issues. Inferring from the theory of critical mass for collective action, we recommend that developers of crowdsourcing tools include a function that allows users to view the contributions of other users.

Keywords: *crowdsourcing, citizen science, information infrastructure, design principles, IS evaluation*

Cite This Article: Selamawit Molla Fossum, Susana Lopez-Aparicio, and Håvard Vika Røen, "An Infrastructural Analysis of a Crowdsourcing Tool for Environmental Research." *American Journal of Information Systems*, vol. 6, no. 1 (2018): 13-22. doi: 10.12691/ajis-6-1-3.

1. Introduction

Defining information infrastructures as an open, shared and heterogeneous socio-technical complex, the existence of two major design problems in the life cycle of information infrastructure development processes were identified [1,2]–these are the bootstrap and the adaptability design problems. The first design problem concerns how to get started or attain a critical mass of participants. The second problem is related to how to sustain the infrastructure through time by adapting to future needs.

Drawing primarily on the first design problem and on the theory of critical mass, we present and analyze a case study that sought to employ crowdsourcing to address an urban environmental challenge. The case study focuses on air pollution, and specifically on emissions from fuelwood used for residential heating in Norwegian cities. The study explored two major crowdsourcing techniques used to collect input data to gain a better understanding of Particulate Matter (PM) emissions. The first technique is conventionally called active crowdsourcing, in which individuals are expected to directly contribute information (e.g. wood consumption) through an adaptation of an existing generic tool or through a new tool developed from scratch by the researchers. In the second method, already existing datasets were assessed if they are useful for emission calculation. The scope of this paper is limited to the active crowdsourcing technique applied by developing a new tool, which from now on is referred to as iR_Wood.

Crowdsourcing as a concept has been defined as an open call for participation in the discovery of knowledge, collection of observation data, generation of new ideas, and the solving of complex tasks [3,4,5]. Crowdsourcing, as the name indicates, places a large number of participants at its core for covering research areas that would be difficult to reach using traditional research methods. The cost-effectiveness and the participatory nature of crowdsourcing has also been the main driving force behind the wider application of the concept in research practices. Its applicability, however, is not without problems. The challenges of engaging participants and doubts about the trustworthiness of the collected data have been reported as primary reasons for not fully incorporating the concept in research practices. In addition, the current crowdsourcing literature does not provide practical guides and evaluation tools. To our knowledge, no prior studies have applied an information infrastructure perspective to understand and propose further strategies as a way to address the user engagement challenge of crowdsourcing.

This paper is, therefore, an attempt to link concepts of information infrastructure with the phenomenon of crowdsourcing. In this sense, crowdsourcing is conceptualized as the making of a part of a larger information infrastructure that needs to be simple and useful for various stakeholders to address the kick-off (bootstrap) problems. Previous information infrastructure bootstrapping studies have been confined to analyzing the issue in organizational settings, where institutional mechanisms are the most prominent means of diffusion. Institutional mechanisms include the use of obligatory regulative mechanisms dictated by concerned administrative units [6]. The case presented in this paper was initiated by researchers, and the crowdsourcing element was expected to operate based on the voluntary participation of individuals, either out of an interest in the environment or because they are concerned about the economic costs of using wood instead of other heating sources (i.e. electricity). Hence, by conducting this study, we intend to contribute to design strategies of crowdsourcing tools by analyzing the diffusion of crowdsourcing tools in the absence of regulative institutional mechanisms. This theoretical understanding will help analyze the practical bootstrap problems of the iR_Wood crowdsourcing tool that is designed in the project the authors are involved. With this background, this study intends to answer the following research question:

What are the socio-technical conditions in crowdsourcing for environmental research that cause a bootstrap design problem?

The rest of the paper is organized as follows. In section two, we present a condensed literature review of crowdsourcing and its application in environmental research. In section three, we describe the data collection methods used to conduct this study. Section four describes the processes of developing a crowdsourcing tool for estimating PM emissions as a result of burning wood. The final two sections are dedicated to the analyses, discussions and conclusions.

2. Crowdsourcing: Overview, Potentials, and Challenges

Crowdsourcing is a concept that takes its blueprint from open source software development ideologies and principles. It is a form of outsourcing in which a portion of a task or business process is delegated to a large online community. The term was first coined by Jeff Howe in Wired Magazine. Howe defines crowdsourcing as “the act of a company or institutions taking a function once performed by employees and outsourcing it to an undefined (generally large) network of people in the form of an open call” [3]. It was first applied in business disciplines, and it has since spread to other fields such as science (e.g. citizen science), urban planning (e.g. Public Participation Geographic Information Systems) and knowledge management (e.g. Wikipedia). The core of the concept is that new ideas, possibilities, and methods of solving complex problems can be found in the crowd that exists outside organizational boundaries. The crowd is characterized primarily as amateur or non-expert in a given field, but collectively it can be smart when individuals’ knowledge is aggregated. Internet-based technologies (including cloud computing & mobile telephones) make it easier to tap into such types of knowledge. One of the most well-known crowdsourcing platforms is the Amazon Mechanical Turk (Amazon.com) in which organizations or individuals outsource small tasks referred to as human intelligence tasks to workforces with Internet access. Tasks in citizen science projects range from passive data collection by citizens to engaging participants in the full cycle of scientific inquiry.

When crowdsourcing is applied to urban environmental research and decision-making, as it was in the study described in this paper, it democratizes scientific research for citizens. For scientists, it broadens the reach of the project and provides a richer data set to the scientific inquiry [7]. In addition, crowdsourcing is an appropriate aid to urban planners. It adds local non-expert knowledge of space designs and citizen insights about the environment that the urban planners might otherwise have neglected [8]. In the field of urban environments, a discipline distinct from urban planning, crowdsourcing has been used as a means of obtaining better knowledge regarding air pollution [9], biodiversity [10], disaster management [11], and climate change monitoring activities [12]. As part of the research project, we identified several crowdsourcing tools for environmental research from literature published between 2006 to 2016. Some examples are illustrated in Table 1.

Table 1. Landscape of crowdsourcing tools for urban environment

Domain	Tools
Urban space design	Innocite, incubators of public spaces, Artmaps, BikeNow, Tutordrive
Air and Noise pollution	EnviroCar, NoiseTube, WideNoise, AirProbe, Mapping for change, Green Dallas, PiMi air box
Disaster management	GeoODK, ebayanihan, MyDisasterDroid, Ushahidi, PataJakarta.org, AsonMaps
Water quality and flood monitoring	HydroCrowd, IDAH2O, Open water, Cyberflood, Floodpatrol, CreekWatch , SimDelta, CrowdHydrology
Biodiversity monitoring	FireflyWatch, Christmas Bird Count, Monarch Larva, BRAND SCAN,Birds calander, BAMONA
Others	MyGeoTrust, CoCoRAHS, Adaptive gazing, Citizenscience.gov, OpenIDEO, Publiclab UK, Open Data Aarhus, Sustainable City Network

Most of the tools are made for an online community and serve for both data collection and visualization. This demonstrates an important aspect of a crowdsourcing tool-data visualization.

The primary challenge of crowdsourcing lies in recruiting, managing, and motivating contributors throughout the entire process [8,13]. Previous studies indicate that the motivation behind engaging in such projects is mainly intrinsic; referring to enjoyment, moral obligation, and the need to protect nature [12,14,15]. Extrinsic motivation factors such rewards of money and other goods have also been recommended by theoretical motivation studies (ibid.). Practically, from the decade of research on crowdsourcing for environmental research, we found only one study that recruited crowd workers for a geofencing task via payment [16]. All other studies encouraged engagement through volunteerism and for altruistic reasons.

The nature and organization of a crowdsourcing project are also important factors; especially in terms of providing feedback and allowing direct contact with coordinators, which play a significant role in recruiting and retaining participants [14]. Furthermore, incorporating gamification in the technologies enables a certain degree of participation from younger participants [17]. People are inherently heterogeneous. For example, some individuals are interested in generating ideas, while others are more interested in giving feedback to the presented ideas. Crowdsourcing applications are recommended to be developed to accommodate this range of involvement and thereby ensure the various groups are reached [5].

In contrast to other crowdsourcing activities like Wikipedia, where contributors are immediate users of others' contributions, in citizen science, there is a significant delay between the time the contribution is made and the time the contribution is made public (i.e. scientific results) [18]. This may pose additional challenges to retaining motivated users throughout the entire process.

The second major challenge of using crowdsourcing in environmental research is associated with concerns regarding the quality of the contributed data. Various measures to address this concern have been suggested based on case studies. The suggestions include proper instruction, training, and the use of controlled variables [12,19,20,21,22].

The nature of crowdsourcing as an online activity with anonymous participants and the relative absence of policy and ethical standards may give rise to unethical practices in citizen science projects. Privacy issues of how contributed data is used and potential surveillance of contributors have also been pointed out as factors having implications for engaging the crowd in the field of research [23]. However, we have not yet identified a research which focuses on the relationship between design and participant engagement. This study contributes to this knowledge gap.

3. Information Infrastructure: Core Concepts

The information infrastructure (I.I.) perspective represents contemporary thinking within information systems (IS) research. Information infrastructure is a means of understanding large, complex, and interconnected IT based systems used in areas that are not in the control of one organization. The growth of such large systems, however, starts from a simple application like the iR_Wood.

Prior works on I.I. have characterized I.I.s as IT capabilities, applications, platforms, and I.I.s in order of increasing complexity [2]. An IT capability is defined as "the possibility or right of the user or a user community to perform a set of actions on a computational object or process" (p.2). Potentially, by following the principle of positive network effects, the value of the IT capability for each user increases when the total number of users increases. When the user base grows, the IT capability increases; and thus, increasing further the user base. This is known as the self-reinforcing effect [24] of the IT capability. In connection with this, the question that is relevant to us, is how to attract these initial users of the iR_Wood tool who can in turn further draw other users. Moreover, there is a need to determine what barriers exist to engaging potential users. Concepts of installed base, cultivation, and bootstrapping are particularly useful in responding to such questions.

Installed base: Strictly speaking, installed base represents the number of installations of software and its users. The emergence of cloud computing and web-based systems means that we cannot base any analysis on counting the number of installations, because one installation can be used by several users. In I.I., an installed base also includes the existing practices of the heterogeneous user base, surrounding institutions (i.e. formal and informal rules) and other existing computer based systems [25]. The installed base encompasses everything that is there prior to the development and use of a new IT capability and it has a determining role in the growth of the IT capability or the application. That is, it either enables or stunts growth at various stages of the I.I. development.

Understanding the potential of the installed base as an enabler or constraint on the use of an application, [26] identified three installed base approaches that were followed by projects. Those are installed base-friendly, installed base-ignorant, and installed base-hostile. From their cross-case analysis, the authors concluded that installed base-friendly approaches that acknowledge and support existing work practices, introduce simple solutions, and are built upon existing technologies that are likely to reach the stage where the adoption and use of the technology is sustainable.

Cultivation: Cultivation is a human-centered design) strategy which gives due attention to the friendly and opposite role of the installed base during the development and adoption of information technologies. As opposed to traditional software development approaches such as the water fall and evolutionary approaches, the cultivation approach recognizes the fact that any software is part of an already existing information technology and associated social practices. Thus, by studying experiences we would be fulfilling the need to learn the installed base. Moreover, the cultivation approach recommends that designers follow small scale incremental changes [25,27] and understand change as a process [28]. The central argument is that the accumulated existing practices, regulative and normative institutions, and technological tools create inertia against the successful uptake of novel innovations and that an ample amount of time should be given to creating appropriate values to targeted users.

Bootstrapping: Bootstrapping is a concept that acknowledges the challenges of getting the first group of users to reach critical mass. To get the first group of users, the bootstrapping approach recommends 1) leaving space for possibilities for action; 2) focusing on the most knowledgeable user groups because those with knowledge of the issue are also the most motivated; and 3) the technology should be simple, flexible, and future-oriented [1]. The technology aspect (the IT capability) is further elaborated in the design principles. The bootstrapping principles present broad guidelines for designers on where to focus and how to shape their design templates. The principles and associated design rules are summarized in Table 2.

Table 2. Bootstrapping design problems, principles, and design rules (source: [2])

Design Problem	Principles	Design rules
Bootstrap: The design goal here is to generate attractors that bootstrap the installed base	Design initially for direct usefulness	DR 1. Target IT capability to a small group DR 2. Make IT capability directly useful without the installed base DR 3. Make IT capability simple to use and implement DR 4. Design for one-to-many IT capability in contrast to all-to-all capabilities
	Draw upon existing installed base	DR 5. Design first IT capability in ways that do not require designing and implementing new support infrastructure DR 6. Deploy existing transport infrastructure DR 7. Deploy build gateway to existing service and application infrastructure DR 8. Build bandwagon associated with other I.I.s
	Expand installed base by persuasive tactics	DR 9. "user before functionality"-grow the user base always before adding functionality DR 10. Enhance any IT capability with n the I.I. only when needed DR 11. Build and assign incentives so that users have real motivation to use the IT capability with in the I.I. in new ways DR 12. Develop support communities and flexible governance strategies for feedback

In the first design principle, the core problem covers how to attract early adopters, who take the risk of adopting the IT capability without benefiting from the network effect of the installed base. Designing the IT capability to be simple and cheap while also being directly useful for those early adopters is expected to generate the first group of adopters.

The second design principle focuses on the importance of the installed base. Designers are advised to build their IT capabilities on applications or platforms that already have a good user base, in order to capitalize on the "bandwagon effect".

The third principle presents the importance of strategically expanding the installed base through incentives and other motivational mechanisms. It also recommends integrating the IT capability with existing platforms and applications. The third principle requires the ability to enroll more users through feedback, incentive, and mutual learning. Previous research has applied the principles fully to studying successful national health information infrastructures [27,29]. We use the principles to study partial failure, as it is equally important for drawing lessons and further nourishing the I.I. growth.

4. The Theory of Critical Mass

Although I.I. identifies bootstrapping as a design problem, and provides helpful guidelines for practice, its applicability has been limited to organizational settings, where top level management decisions play a vital role in the bootstrapping processes. Our case deals with collective action; engaging urban citizens for the collective goal of improving air quality in the absence of a regulatory authority. Hence, we draw on [30,31] the theory of critical mass to complement bootstrap design principles. Based on the theory of collective action, [30,31] argue that "collective action usually entails the development of a critical mass-a small segment of the population that chooses to make big contributions to the collective action while the majority do little or nothing" (p. 524). Unlike other collective action theorists, who assume individuals make isolated and independent rational decisions about contributing, [31] identify interdependency in decisions as a driver for collective actions. They assume that "individuals take account of how much others have already contributed in making decisions about contributing to a collective action" (p.524). They, however, do not explicitly state how individuals determine the contributions of others; something information technologies can facilitate. Decision interdependency follows the logic of reciprocity. It means when individuals see contributions made by others, they become willing to contribute their fair share to securing a collective good for altruistic reasons and like dispositions without the inducement of material incentives [32]. The logic of reciprocity views individuals as reciprocators and the more individuals contribute, the more others are willing to contribute; leading to a state of self-reinforcement [24].

In addition to interdependency, we find their concept of group heterogeneity pertinent to our analysis of crowdsourcing as a tool for improving the understanding of air pollution in the urban regions of Norway. There usually is a range of interests or desires in groups regarding a collective good. For example, "although we all want clean air, those among us suffering from emphysema want it more. For those with homes of equal value, a potential school closing is more important to those homeowners with school-age children than to those without" [[31], p.528]. This therefore requires the identification of the different groups and targeting communication for each group, shaping the issue to appeal to their interests.

These concepts were used as methodological guide during data collection and as a lens for analyzing the data.

5. Data Collection Methods

This study followed a qualitative case study research method as described by [33]. Yin underscores the need for a thorough literature review before designing a case study. Accordingly, we conducted a systematic literature review of crowdsourcing for environmental research (ranging from 2006-September 2015) to identify the main challenges and opportunities of crowdsourcing for urban environmental studies [34]. This case study builds on those findings; and applies some of the strategies to understanding the enabling and constraining factors to bootstrap the newly developed wood consumption web application. It also helped us design the research. First, we interviewed core participants of the project to identify their practical challenges. The responses from these participants then led to interviews of key stakeholders regarding their views on the tool, crowdsourcing, and whether or not the tool would be useful for them. The literature review was specifically used to identify available crowdsourcing tools for environmental research and to identify user engagement mechanisms.

The unit of analysis in the case study is the air pollution sub-case in the iReponse project (http://iresponse-rri.com/). iResponse is a research project consisting of three research institutes, two universities, a designing company, software as a service (Saas) company, and a group of key stakeholders. All of the research institutes are located in Oslo, Norway. The software company and one of the universities are located in Helsinki, Finland. The project is funded by the Research Council of Norway and aims to explore socially-responsible crowdsourcing methods and tools for urban environmental research and decision-making. It follows principles of responsible research and innovation [35]. In addition to air pollution, the project is engaged in applying crowdsourcing to storm water management and urban space planning.

Data were collected using qualitative data collection techniques such as interviews, document analysis, and participant observations. We interviewed project partners and key stakeholders to get their opinions regarding crowdsourcing as a method for urban environment research. As the project partners and stakeholders were found to have different roles and interests in the project, we adapted the interview questions taking into consideration their individual roles and institutional interests into account. For example, while some of the participants were more concerned about how to protect citizens' interests in crowdsourcing initiatives, others were more involved in designing the crowdsourcing tool in socially-responsible and user-friendly ways. Most interviews were recorded after obtaining verbal consent from the respondents and those interviews were fully transcribed. For those interviews that were not recorded, we sent interview notes back to the interviewees to validate the accuracy of the notes. In total, we interviewed 18 individuals. The following Table summarizes the characteristics and number of individuals interviewed.

Table 3. Characteristics and Number of Informants

Type of respondent	Number
Core participants	2
Other project participants	3
Designers	2
Key stakeholders	11
Total	18

In addition to interviews, project documentation produced in the past year and a half has served as a data source to build up the case description presented in the following section. The referenced project documentation is summarized in Table 4.

Table 4. Project documentation used as data source

Document type	Year
Minutes of project meetings	October 2015-March 2017
Stakeholder workshops	October 2015 & October 2016
Design sketches and presentations	October 2015-September 2016
Citizen survey by consumer research institute [36]	Sept. 2016
Air case study experience and lessons learned (air research institute)	27th February 2017
Emails correspondences	Sept 2016-

6. Case Description: Crowdsourcing Input Data for Woodburning Emissions

6.1. Background

In Norway, there is a historically embedded tradition of using firewood for heating in the winter season. The Norwegian National TV (NRK) program head was quoted by BBC saying, "People in Norway have a spiritual relationship with fire.... fire is the reason we're here, if there was no firewood, we couldn't live in Norway, we'd freeze" [37]. Wood is not burned only for heating purpose. Fireplaces also serve a social function, providing a gathering point for friends and family [38]. On the other hand, research identifies fireplaces and wood stoves as generators of ambient particles smaller than 2.5 microns in diameter (PM 2.5). These particles are produced by incomplete combustion regimes. Studies have linked particles of such kind with negative health and environmental effects, albeit disagreement exists on the health effects due to questions regarding the methodologies. In general, air pollution caused by wood smoke has been less studied than pollution generated by vehicle traffic, which makes designing control measures a challenge. The common approach to minimizing its adverse effects has been to promote responsible wood-burning through mass media. For instance, in February 2017, the city government of Oslo launched a campaign describing how to start a fire correctly for safety and environmental reasons. One of the most important measures is the implementation of economic incentives for shifting from older to newer appliances. This incentive has existed in the municipality of Oslo since 1998. To measure the impact of such campaigns and propose new solutions to the pollution problem, there is a need for up-to-date geolocalized input data on, for instance, wood consumption and type of stove technology. The currently-available data is insufficient to properly determine the impact of burning wood on ambient air pollution, particularly in urban areas. A respondent explained the situation as:

[Data] is available at the regional level, but they are not at the level of detail needed to address urban issues. They are aggregated. They are not in high spatiotemporal resolution. We need geographic details. The data should also be up-to-date. Weather conditions change, so we should collect new data at least once every two years.

PM emissions for this purpose are calculated as follows:

$$Emissions_{woodburning} = \sum Wood\ Consumption_{type\ of\ oven} \times emission\ factor_{Type\ of\ oven}$$

For estimation of emissions at high spatial resolutions, input data concerning urban residents' frequency of wood use, the types of stoves used (i.e. open fireplace, closed stove), and years the stove were produced (i.e. closed stoves produced before or after 1998), and the location of the householder are needed.

Compared to using sensors, crowdsourcing is considered a cost-effective, participatory, and minimally invasive method of testing and collecting input data. It has been tested in two ways. The first one is using an existing crowdsourcing platform, as reported in [39]. The second one, which is the focus of this research, is by developing and testing a new tool using a participatory approach.

6.2. Current Wood Consumption Data Collection Practice

Currently, data on wood consumption are collected by Statistics Norway as part of the quarterly holiday and travel survey. The survey gathers data that covers the preceding 12 months' wood consumption. The question about wood consumption is carried out three times a year, with a total population size of 6,000 persons per year. Over the course of a year, a total net sample of almost 3,200 are carried out. The figure used in the emission calculation is the average of the three quarterly surveys. Statistics Norway's population database is used to define the surveyed population and to select samples. Each survey is based on a representative sample of 2000 persons between the ages of 16-74 and the response rate averages 60 percent.

Data collection is primarily carried out using a telephone interview (computer assisted interviews). There are several procedures for electronically controlling the registration of answers. Based on this data, Statistics Norway estimates the amount of wood used in Norway every year and calculates the emissions of the burned wood. Emissions data are annually provided to the Norwegian Environmental Agency and reported to the Economic Commission for Europe: Conventions on Long-Range Transboundary Air Pollution (CLRTAP). The statistics are to a large extent compiled to cover the demands of reporting to CLRTAP on behalf of the Ministry of Climate and Environment. The data is collected at the national level and it is break down to county and municipality levels. Map-based data by municipality are not available. The following picture depicts the existing top-down data collection practices and the crowdsourcing bottom-up data collection structure we recommended

Figure 1. Top-down Vs bottom-up approach

For introducing crowdsourcing, we developed iR_Wood over 14 months following a participatory approach. We involved stakeholders and project participants in defining the structure and design of the tool, as the variables that need to be collected are very much constrained by emission estimates. Moreover, the design was open to citizens' feedback regarding the user-friendliness of the tool. The first plan involved the development of a mobile app which would collect the names and addresses of residents. Those variables were removed for privacy reasons, and variables that are useful to scientists but do not expose the identity of residents were recommended. For example, instead of names, postal address was recommended for establishing proxies as sufficient or estimating emissions due to fuel wood consumption.

The iR_Wood tool was designed to provide feedback to citizens, which was meant to motivate them to report their wood consumption practices on a regular basis. Upon entering the last input data, which is the postal code, the participant gets individualized feedback in the form of a comparison of the costs of burning wood with costs of using electricity for heating (Figure 2).

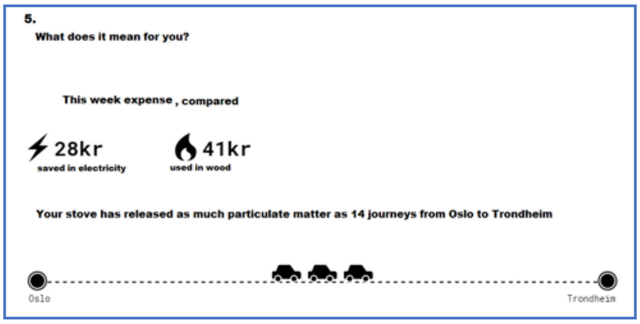

Figure 2. Individualized feedback

Furthermore, the iR_Wood tool gives a personalized estimate of the amount of PM released from a household's reported wood consumption, comparing it with the PM emitted by cars (Figure 2). This is meant to make individuals aware of their contribution to air pollution in urban areas. However, the tool, at this point, does not provide feedback on cumulative economic loss by aggregating previous reports, nor does the tool provide a means of comparing individual wood consumption behaviors. We assume that adding such a comparative in the system would activate self-reinforcing mechanisms.

After receiving their feedback, wood consumers are asked about two more variables: the time of the day when they use their stoves (e.g. morning, afternoon, evening) split in weekdays and weekend, and the type of dwelling in which the reporter lives. All of these variables are of vital importance in establishing variations of wood-burning activities over time. Toward the end, the participants are given the opportunity to contribute with ideas about how to reduce the air pollution associated with burning wood. Among the ideas collected by citizens, technological and regulative measures such as additional economic incentives to shift from old to newer efficient stoves, or even to ban burning wood during high pollution episodes are highlighted.

The iR_Wood crowdsourcing tool is intended to be integrated with national environmental monitoring platforms, once users start to use it.

6.3. Implementation

The iR_Wood tool was operational during a short winter season (January – March 2017). Different approaches were used to attract citizens to the tool and draw their attention to the problem of air pollution caused by burning wood. These include advertising through Facebook, promotion on Twitter, distribution of postcards, and the use of personal and professional networks. Dissemination strategies such as television, newspapers and online marketing were not used, and have been mentioned by several interviewed participants as instrumental to reaching a wider range of the population. Respondents indicated that this, in addition to the use of different communication media including Google AdWords and the involvement of government agencies would result in better public engagement in crowdsourcing projects.

6.4. Post Implementation Reflections

By the end of April 2017, around 72 participants had used the tool to report their wood consumption. The participants were from different Norwegian urban areas. Specifically, 29 of them were from Oslo, four from Bodø, and the remainder spread across different cities. The researchers noted the low participation and concluded that the collected data could not be used for research purposes. The core user group and stakeholders provided their views regarding the challenges and possible alternative strategies. A respondent explained the situation as follows:

The tool is not working for the purpose that it was designed, as there is a very low level of participation. In my opinion, the reason is that it requires people to report regularly, on a weekly basis, and people are busy. We cannot generalize from the limited number of participants, as it is not a representative sample. In my experience, it is better to ask people to report only once about the entire winter season. The tool is incompatible with lifestyle of the society. I believe it could be a good idea if it is kicked off by authorities and not by scientists.

Another respondent stated the user engagement problem as follows:

The challenge is in the tool itself. I do not think anybody is interested in using it. You need to login and so on. So it is not used. We have been discussing how it should be used and maybe we should use schools. We should make the young generation interested. We have been discussing this, but we did not try. It was an idea as engagement is the major problem in our work. From the development side, it was really smooth. It is the user side which was the main challenge.

Some respondents did not seem very eager about reporting wood consumption, and they were not surprised by the low response rate of the tool. One respondent, for example, challenged the very idea of wood consumption as a cause of air pollution and how the issue itself is not interesting:

First, the idea of reporting wood consumption can be a boring subject. At least if it is wood, it is already cut and whether you leave it or burn it, you will have the same CO_2 emissions. In this sense, burning wood is supported in climate change debates as climate friendly [this quote is rephrased].

Taking into account that it takes time to design innovative products, others consider the low participation rate to be a typical problem. In the eyes of some, all that is required is more time and the use of different experimental designs and feedback mechanisms. A respondent stated:

It is not unexpected that it is not used. It is perhaps a bit different views about the scope of the problem they are trying to solve. But there are a lot of start-ups that try to make a product. To go from idea to actually reaching users is really hard work and takes a lot of time. I would expect 15 tries in different pages and distribute it. I do not see it as a failure. It is just expected, try and fail before making it a success.

Others indicated the need for regular reminders by email or SMS, because otherwise they may forget to report their wood consumption regularly.

6.5. Usefulness of the Application

Key stakeholders were asked if they would be interested in collecting data using a crowdsourcing method. From their responses, it was evident that although they want to know the distribution patterns and concentrations of wood smoke more accurately to allow them to assess fire risk and measure the impact of wood smoke on health and the environment; they also expressed doubt regarding crowdsourcing for emission inventory. For example, respondents explained that the crowdsourcing approach could help differentiate between the different sources of PM and enable them to identify sources that contribute the most to health effects.

The stakeholders were especially concerned with the question of whether or not the data accurately represent the whole population, and the need to know the wood-burning practices of individuals. The stakeholders repeatedly mentioned that data collected using the current crowdsourcing approach does not follow typical statistical methods. The data is collected from people with Internet access and with knowledge of the impact of wood smoke. It, therefore, does not account for different population groups, thus making crowdsourcing a less interesting approach for them. A respondent who is working in the field of emission inventory expressed his view as follows:

Crowdsourcing could be interesting for getting more accurate data on technology used and the amount of wood used in different technology but there is a time and money issue. It seems very demanding to get people report their wood consumption, but how do you ensure that your sample is representative?

Wood-burning practices are not part of the online questionnaire in iR-wood. It is expressed as an important variable in understanding emissions. The combustion regime has equal impact on how much soot is emitted. Therefore, without this data type, the tool becomes less useful. As a result, stakeholders recommended the use of measuring stations and chemical analyses to determine whether the particles are from wood or from other sources rather than trying to acquire data directly from residents; which they assumed would be difficult and boring.

7. Analysis and Discussions

In this section, we will analyze the case described in the previous section according to bootstrapping design principle [1,2,26]. We will also use the concepts of interdependency and group heterogeneity as presented by [30,31].

7.1. Design Principle 1: Design for Direct Usefulness

The crowdsourcing tool was initially designed to be immediately useful to scientists. Scientists needed high resolution data regarding the wood consumption behavior of urban residents to better understand the contribution of wood smoke to ambient air pollution. The existing data sources were found to be insufficient to properly estimate PM emissions at the scale of single cities. In addition, solving the current environmental problems need the engagement of society; making a crowdsourcing approach important.

The IT capability of the crowdsourcing tool was primarily designed to address the needs of scientists. The tool was designed to inform scientists of the types of fire appliance, amounts of wood consumed, when wood was consumed, and the locations of wood stoves; all of which are required to estimate PM emission levels.

The IT capability is also targeted to benefit urban residents by enabling them to recognize the economic cost of using wood as compared to using electricity as a heating source. The IT capability also enables urban residents to learn of the PM emissions from their wood stove. To communicate this message in a simple format, designers chose to use the well-known air pollutant measure of transportation emissions as a comparison to wood stove generated emissions (Figure 2). In this way, IT capability was designed to the direct usefulness of scientists and urban residents who use firewood, independent of the installed base. However, the low participation from the wood consumer side and the effort to introduce the tool from the scientist's side indicates that iR_Wood is more useful to the scientist than to the wood consumer.

In terms of the recommendation to make IT capability simple to use and implement, from the perspective of functionality, the tool is simple to use and implement. It resembles a survey tool in which users respond to questions, with no additional training or extended instruction. However, the overall aim of reporting wood consumption was not well-received. In addition, users who continuously report need to log in so they would not need to report some of the variables such as the type of wood stove. In comparison to other crowdsourcing tools that have been tested in this project, which do not require login, the need to login with an email address and password is a possible hindrance to user engagement.

In the process of making the tool useful for early adopters, there has been no identification of different user groups. The process assumes homogenous groups of firewood users and expects all to volunteer to report their wood use in a continuous manner. The aim of this is to draw collective actions for the common good—good air quality. Crowdsourcing projects are primarily dependent on volunteers, and consistent with the critical mass theory, there is a need to recruit volunteers or paid workers to bootstrap the crowdsourcing initiative for collecting wood burning data.

7.2. Design Principle 2: Draw Upon Existing Installed Bases

The installed base in this case encompasses the existing data collection practices for wood consumption, the existing IT infrastructure, and the deeply embedded cultural value of burning wood; not necessarily for heating but as a social function.

Given the existing data collection practices, crowdsourcing can be seen as a disruptive approach, in direct opposition to the installed base. While data on wood consumption is currently collected using what can be referred to as a top-down approach, crowdsourcing promotes a bottom-up approach based on data collection from individual users and enabling aggregation of data by city or county as well as on the national level. The positive cultural value of heating with wood may also partially explain the low participation rate, because the tool exposes users' wood usage behavior. In addition, there are issues of whether the data collected using the existing practices is representative enough to support general conclusions about the whole population. These unresolved issues mean that it is difficult to draw on the existing installed base; limiting the ability to benefit from the bandwagon effect generated by the installed base. Based on the characterization of different information systems projects relative to the installed base provided by [26], we found the current crowdsourcing approach installed base-ignorant, which might have led to the participation problem. Hence there is a need to shift to a cultivation approach in which the installed base is taken into consideration by focusing on understanding the needs of key stakeholders.

In terms of existing IT infrastructure, the iR_Wood tool was developed without needing any additional support—at least from the end user's side. Once it gets started, it is intended to be integrated mainly as a data source for the national environmental monitoring platform (e.g. http://www.miljostatus.no/; http://www.luftkvalitet-nbv.no/). This requires expanding the install base using persuasive tactics.

7.3. Design Principle 3: Using Persuasive Tactics to Expand the Installed Base

After putting the iR_Wood tool into use, different campaigning strategies were employed, mainly targeting firewood users. The campaign strategies yielded little input from the primary target users. We then explored how the tool could be useful to other users who need wood consumption data directly or indirectly. Following the principle of "users before functionality", we presented key stakeholders with the potential uses of the tool for their work. They expressed interest in cooperating as the project offered additional benefits to their work. This strategy of persuading key stakeholders should continue in further development phases of such tools to fulfill some of their data needs. For example, data variables that reflect the wood-burning practices need to be added to satisfy the data needs of key stakeholders.

One of the aims of the iR_Wood crowdsourcing tool is to mobilize collective action for managing environmental challenges. In this case, as informed by [31], individual decisions to contribute for the collective good depend on the perception of contributions of others. This then is a reminder to design crowdsourcing tools to support reciprocity of action. This could be achieved by incorporating functionalities that compare each user's cumulative use of wood with that of others. From literature, we observed aggregate data visualization as a key characteristics of crowdsourcing tools.

8. Conclusions

This paper is based upon a study that sought to crowdsource geolocalized data on wood consumption and stove types for research and decision-making purposes. One of the key aims was to design and implement a simple tool to address the needs of scientists. Despite the effort to draw wood users to the tool, only a small number of participants registered. By analyzing the case using concepts drawn from information infrastructure and the theory of critical mass, we revealed the socio-technical conditions that need to be fulfilled to successfully bootstrap the tool.

From the technical side, our analysis mainly illustrates that crowdsourcing tools designed to collect and report research data should support reciprocity of action. One way of achieving reciprocity is to reveal the cumulative usage of individual users as compared to other users. From the social side, we identified three groups of users to whom iR_wood needs to be useful in order to gain momentum. These are scientists, wood consumers, and key stakeholders. While the tool as currently designed is directly useful to the scientists, there are doubts of its usefulness to wood consumers. For the tool to be useful to wood consumers, designers must recognize the heterogeneity of wood consumers, and a small number of early adopters should be recruited using extrinsic motivation strategies. To make the tool directly useful to stakeholders, issues of data representativeness and wood-burning practices need to be addressed. This means that the first step of designing a crowdsourcing tool for environmental research should be making the tool useful for various stakeholders who require the same information for varying purposes. For example, the fire department is interested in wood consumption data for risk assessment. Hence, making the tool useful for them would expand the user base, as reporting to reduce fire risk might motivate more users than reporting for research.

Lastly, we found that information infrastructure concepts useful when designing and managing a crowdsourcing tool. I.I. provides principles and rules that help expose and address the challenges of participant engagement in crowdsourcing projects. However, the bootstrap principles should be modified to include decision interdependency to better suit the nature of crowdsourcing platforms for research, which is creating online citizen scientists. Hence, we recommend one additional design rule. DR13: An IT capability should be built to allow users to view the contributions of others.

Acknowledgements

This study was made possible with the financial support of the Research Council of Norway (247884/O70) through the financing of the iResponse project. We thank the participants in the iResponse project for their support and useful feedback. Special thanks to Gjermund Alsos and Will Hindson from Netlife Research for designing the tool, and to Matthias Vogt for his continuous support in carrying out the case study. We would like to thank Kristin Aasetad from Statistics Norway for providing data on the current data collection practices.

References

[1] Hanseth, O. and M. Aanestad, *Design as bootstrapping. On the evolution of ICT networks in health care.* Methods of information in medicine, 2003. 42(4): p. 384-391.

[2] Hanseth, O. and K. Lyytinen, *Design theory for dynamic complexity in information infrastructures: the case of building internet.* Journal of information technology, 2010. 25(1): p. 1-19.

[3] Howe, J., *Crowdsourcing: How the power of the crowd is driving the future of business.* 2008: Random House.

[4] Estellés-Arolas, E. and F. González-Ladrón-De-Guevara, *Towards an integrated crowdsourcing definition.* Journal of Information science, 2012. 38(2): p. 189-200.

[5] Brabham, D.C., *Crowdsourcing*. 2013: Wiley Online Library.

[6] Fossum, S.M., *An institutional perspective on health information systems' standardization: multiple case studies*, in *Department of Informatics*. 2016, University of Oslo: Oslo.

[7] Augar, N. and M. Fluker. *Developing social media for community based environmental monitoring*. 2014. ACIS.

[8] Brabham, D.C., *Crowdsourcing the public participation process for planning projects.* Planning Theory, 2009. 8(3): p. 242-262.

[9] Henriques, J.J., et al. *Crowd-sourcing urban air-quality in developing countries through open source technologies*. in *Global Humanitarian Technology Conference (GHTC), 2013 IEEE*. 2013. IEEE.

[10] Heaton, L., et al., *Participatory Science: Encouraging public engagement in ONEM.* International Journal of Science Education, Part B, 2016. 6(1): p. 1-22.

[11] Shahid, A.R. and A. Elbanna. *The Impact of Crowdsourcing on Organisational Practices: The Case of Crowdmapping.* in *ECIS*. 2015.

[12] Gharesifard, M. and U. Wehn, *To share or not to share: Drivers and barriers for sharing data via online amateur weather networks.* Journal of Hydrology, 2016. 535: p. 181-190.

[13] Fröhlich, S., et al. *BikeNow: a pervasive application for crowdsourcing bicycle traffic data*. in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 2016. ACM.

[14] Baruch, A., A. May, and D. Yu, *The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform.* Computers in Human Behavior, 2016. 64: p. 923-931.

[15] Aitamurto, T., H. Landemore, and J. Saldivar Galli, *Unmasking the crowd: participants' motivation factors, expectations, and profile in a crowdsourced law reform.* Information, Communication & Society, 2017. 20(8): p. 1239-1260.

[16] Konomi, S.i. and T. Sasao. *Crowd geofencing*. in *Proceedings of the Second International Conference on IoT in Urban Space*. 2016. ACM.

[17] Brovelli, M.A., M. Minghini, and G. Zamboni, *Public participation in GIS via mobile applications.* ISPRS Journal of Photogrammetry and Remote Sensing, 2016. 114: p. 306-315.

[18] Antelio, M., et al. *Qualitocracy: A data quality collaborative framework applied to citizen science*. in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*. 2012. IEEE.

[19] Hassanzadeh, R. and Z. Nedovic-Budic, *Identification of earthquake disaster hot spots with crowd sourced data*, in

Intelligent Systems for Crisis Management. 2013, Springer. p. 97-119.

[20] Davami, E. and G. Sukthankar. *Evaluating trust-based fusion models for participatory sensing applications*. in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 2014. International Foundation for Autonomous Agents and Multiagent Systems.

[21] Degrossi, L.C., et al. *Flood Citizen Observatory: a crowdsourcing-based approach for flood risk management in Brazil*. in *SEKE*. 2014.

[22] Fonte, C., et al., *VGI quality control.* ISPRS Geospatial week 2015, 2015: p. 317-324.

[23] Chiu, C.-M., T.-P. Liang, and E. Turban, *What can crowdsourcing do for decision support?* Decision Support Systems, 2014. 65: p. 40-49.

[24] Grindley, P., *Standards, strategy, and policy.* 1995.

[25] Hanseth, O. and E. Monteiro, *Understanding information infrastructure.* Unpublished Manuscript, Retrieved on 6th September from http://heim. ifi. uio. no/~ oleha/Publications/bok. pdf, 1998.

[26] Aanestad, M., et al., *Information Infrastructures and the Challenge of the Installed Base*, in *Information Infrastructures within European Health Care*. 2017, Springer. p. 25-33.

[27] Aanestad, M. and T.B. Jensen, *Building nation-wide information infrastructures in healthcare through modular implementation strategies.* The Journal of Strategic Information Systems, 2011. 20(2): p. 161-176.

[28] Dahlbom, B. and L. Mathiassen, *Computers in context: The philosophy and practice of systems design*. 1993: Blackwell Publishers, Inc.

[29] Jensen, T.B., *Design principles for achieving integrated healthcare information systems.* Health informatics journal, 2013. 19(1): p. 29-45.

[30] Marwell, G. and P. Oliver, *The critical mass in collective action*. 1993: Cambridge University Press.

[31] Oliver, P., G. Marwell, and R. Teixeira, *A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action.* American journal of Sociology, 1985. 91(3): p. 522-556.

[32] Kahan, D.M., *The logic of reciprocity: Trust, collective action, and law.* Michigan Law Review, 2003. 102(1): p. 71-103.

[33] Yin, R., *Case Study Research: Design and Methods, SAGE Publications.* California, USA, 2009.

[34] Fossum, S.M., *Potentials and challenges of using crowdsourcing for urban environmental sustainability*, in *NTNU Sustainability science conference: transitions to sustainable systems*. 2017, NTNU Norwegian University pf Science and Technology: Trondheim, Norway.

[35] Von Schomberg, R., *A vision of responsible innovation.* 2013.

[36] Strandbakken, P., R. Larvik, and H. Throne-Holst, *iResponse: Crowdsourcing; An Approach to Urban Environmental Governance*. 2017, Consumption Research Norway: Oslo, Norway.

[37] BBC. *Norway's NRK broadcasts 12 hours wood buring program.* 2013.

[38] Levander, T. and S. Bodin, *Controlling Emissions from Wood Burning: Legislation and Regulations in Nordic Countries to Control Emissions from Residential Wood Burning An examination of Past Experience*. 2014: Nordic Council of Ministers.

[39] López-Aparicio, S., et al., *Public participation GIS for improving wood burning emissions from residential heating and urban environmental management.* Journal of environmental management, 2017. 191: p. 179-188.