Architecture of Mobile Crowdsourcing Systems

Frank Fuchs-Kittowski[1,2], Daniel Faust[2]

[1]HTW Berlin
{frank.fuchs-kittowski}@htw-berlin.de
[2]Fraunhofer FOKUS
{frank.fuchs-kittowski, daniel.faust}@fokus.fraunhofer.de

Abstract. This paper proposes a general architecture and a classification scheme for mobile crowdsourcing systems, which are illustrated by two example applications. The aim is to gain a better understanding of typical functionalities and design aspects to be considered during development and evaluation of such collaborative systems.

Keywords: crowdsourcing, crowdsourcing system, mobile crowdsourcing, crowdsourcing application, architecture, classification scheme.

1 Introduction

Many organizations are increasingly using crowdsourcing as a new model for value creation, where new web technologies are used to outsource tasks, which are traditionally performed by a specialist or a small group of experts, to an undefined large group of people [1]. Meanwhile, mobile devices (phones, smartphones, tablets, and in the near future glasses, watches, and so on) have become ubiquitous and a tool for crowdsourcing [2]. In mobile crowdsourcing mobile devices are used for data-collection tasks delegated to a larger number of people as well as for the coordination among the people involved.

In the recent years numerous mobile crowdsourcing applications have been realized and have shown the potential for business and society [3]. As the popularity of these applications increases, our understanding of how to design and deploy successful mobile crowdsourcing systems must improve [4]. Many systems described in the scientific literature were individual, task-specific, ad-hoc implementations [5]. Without a profound theoretical foundation, the development of such mobile crowdsourcing applications is still a difficult task and costs as well as the time needed for each development can be high.

The objective of this paper is to gain a better understanding of typical functionalities and design aspects to be considered during development and evaluation of such collaborative systems. Thus, the way how mobile crowdsourcing applications are developed will shift from an ad-hoc manner to a planned routine. Based upon an extensive literature review, a categorization of existing applications of mobile crowdsourcing systems, and an overview of typical design aspects of mobile

adfa, p. 1, 2011.

crowdsourcing systems, a classification scheme and a general architecture for mobile crowdsourcing systems are described.

The remainder of the paper is structured as follows: The second section gives an overview of related conceptual work in the domain of mobile crowdsourcing. A classification scheme for mobile crowdsourcing applications is presented in the subsequent section. In section four, a general architecture for mobile crowdsourcing systems is proposed. Based on this architecture and classification scheme two example applications are illustrated in section five. Finally, main conclusions are drawn and further research tasks in this field are identified.

2 Related Work

Our understanding of the notion ‘mobile crowdsourcing’ is that a group of people voluntarily collects and shares data and information using widely available mobile devices (smartphones etc.), where this data is processed and provided via a data-sharing infrastructure to third parties interested in integrating and remixing this data.

Similar approaches are Volunteered Geographic Information (VGI [6]), Public Participatory GIS (PPGIS [7]), Participatory Sensing (PS [8]). But while all these terms are close, they emphasize different aspects. VGI focuses on presenting the captured geospatial data on maps. PS emphasizes the use of the built-in sensors of mobile devices for data capturing. PPGIS focuses on storage and processing of data in GIS. All approaches have in common the volunteer and participatory nature of the data capturing and sharing process.

Over the last several years crowdsourcing has been a vital research field [9].There has been much research regarding (i) the *definition* of crowdsourcing (e.g. [10], [11]), (ii) the characterization of the crowdsourcing *process* (e.g. [12]), (iii) the development of a crowdsourcing *taxonomy* (e.g. [13, 14, 15, 16, 17]), and (iv) the introduction of a *conceptual framework* that supports the designing of crowdsourcing systems (e.g. [18, 19]). However, only little research was conducted to define a crowd sourcing system and its *technical design* precisely [20].

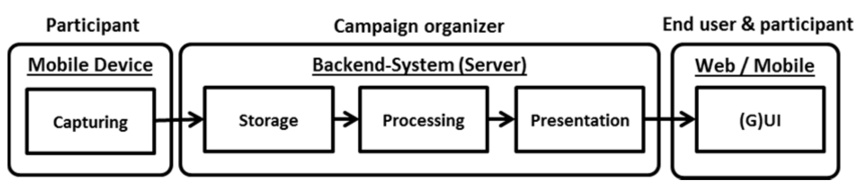

Figure 1. General architecture for Crowdsourcing

In the scientific literature mobile crowdsourcing applications have two essential subsystems/components (see Figure 1): data capturing and data processing [21]. These main components are generally organized as client-server architectures with a mobile client (mobile device) for ubiquitous data capturing, and a server (backend system) for data storage, processing, and visualization [22].

Existing approaches to general architectures focus either on more refined functions of certain subsystems (data capture, data processing, campaign management etc.) or non-functional aspects (privacy etc.). [21], [23] present common architectural components with a special focus on data capturing and leveraged processing. [5] emphasize the functional components regarding campaign management. The architecture presented by [24] pays special attention on recruitment of participants. Several architectures of applications emphasize task management, e.g. the distribution of data capturing tasks and software to the participants resp. their mobile devices [25, 26, 27, 28]. Enhancements of general architectures to solve privacy issues are presented in [29, 30]. Medusa [31], MoCoMapps [32], PRISM [27], and AnonySense [33] are frameworks for cost-efficient development of mobile crowdsourcing applications and with focus on generality, security, scalability, and privacy. A detailed view on the common functional components (and their interfaces) of mobile crowdsourcing applications still does not exist.

3 Research Method

A systematic literature review (SLR) was conducted to improve the understanding on functional and technical aspects of mobile crowdsourcing systems. A SLR provides a well-structured and repeatable procedure to identify, evaluate and interpret existing literature relevant to a specific research question [34]. The main goal of a SLR is not only to methodically aggregate scientific studies in a certain research domain but also to support the development of evidence-based guidelines for practitioners [35].

After formulating the research question (What type of components and functions of mobile crowdsourcing systems can be conceptualized?) an appropriate search strategy was derived and applied: As search resources web search engines (e.g. google search) as well as scientific literature databases such as ACM Digital Library, IEEE Xplore Digital Library, ScienceDirect, SpringerLink and Wiley were used. Any publication type (from application web site to peer-reviewed journal paper etc.) was considered with a restriction to German or English language publications. The literature review includes contributions that (i) describe a certain mobile crowdsourcing application, (ii) address design issues, or (iii) classify or give an overview of applications.

The authors identified 124 publications, which were carefully read. Keywords were collected which either addressed a component or a function of a mobile crowdsourcing application. Iteratively, specific keywords were aggregated to more generic terms. Finally, a concept map was created that maps all relevant literature to one or more of the derived generic components and function terms.

4 Characteristics of Mobile Crowdsourcing Applications

Regarding business cases and application domains, the scientific literature shows a large variety of possible mobile crowdsourcing applications. Based upon the observed object, these applications can be categorized as either people-centric or environment-centric [29]. People-centric applications collect data about the user. They mainly focus on observation of the user's health (e.g., physical effort), documenting activities (e.g., sport experiences), and understanding the behavior of individuals (e.g., eating disorders). Based on this data, the users can be provided real-time support, e.g. in case of an asthma or heart attack. In contrast, environment-centric apps capture information about the surroundings of the user, i.e. environmental parameters (e.g., air quality, air pressure, noise pollution, thermal, road condition, and traffic in cities) or interesting events (e.g., accidents, damages, disasters). Typical application domains are maintenance of man-mad infrastructures (e.g., report of damages in urban areas or observation of depots to reduce costs of maintenance and inspections), environment and nature protection (e.g., mapping of species to take targeted protective measures, monitoring specific geographic areas, and mapping of damages), disaster management (e.g. mapping of damages after natural disaster to use limited resources more efficiently, collecting of real-time-data about a certain area to distribute relief units).

Table 1: Morphological analysis for classification of mobile crowdsourcing applications

<table>
<tr><th></th><th>Characteristic</th><th colspan="6">Feature</th></tr>
<tr><td rowspan="2">Device</td><td>Type</td><td colspan="3">Special device</td><td colspan="3">Standard device</td></tr>
<tr><td>Sensor technology</td><td colspan="2">Manual input</td><td colspan="2">Embedded sensors</td><td colspan="2">External Sensors</td></tr>
<tr><td rowspan="5">Data</td><td>Capturing</td><td colspan="2">Automatic</td><td colspan="2">Manual</td><td colspan="2">Context-related</td></tr>
<tr><td>Spatial data</td><td colspan="2">Point</td><td colspan="2">Line</td><td colspan="2">Polygon</td></tr>
<tr><td>Transmission</td><td colspan="3">Real-time</td><td colspan="3">Delayed</td></tr>
<tr><td>Anonymization</td><td colspan="2">None</td><td colspan="2">Anonymized</td><td colspan="2">Authenticated</td></tr>
<tr><td>Processing</td><td colspan="2">None (raw data)</td><td colspan="2">Analysis</td><td colspan="2">External systems</td></tr>
<tr><td rowspan="6">Participation</td><td>Admission</td><td colspan="2">Own initiative</td><td colspan="2">Non-binding enquiry</td><td colspan="2">Binding request</td></tr>
<tr><td>Selection</td><td colspan="3">A priori</td><td colspan="3">Dynamically</td></tr>
<tr><td>Knowledge/Skills</td><td colspan="2">Low/none</td><td colspan="2">Medium</td><td colspan="2">High</td></tr>
<tr><td>Selection Criteria</td><td colspan="2">Person</td><td colspan="2">Role</td><td colspan="2">Location</td></tr>
<tr><td>Registration</td><td colspan="2">Anonymous</td><td colspan="2">Known</td><td colspan="2">Formal relation</td></tr>
<tr><td>Assessment</td><td colspan="2">Automatically</td><td colspan="2">Participant</td><td colspan="2">Organizer</td></tr>
<tr><td rowspan="4">Involvement</td><td>Degree</td><td colspan="2">Active</td><td colspan="2">Limited</td><td colspan="2">Passive</td></tr>
<tr><td>Task type</td><td colspan="2">Initial collection</td><td colspan="2">Update</td><td colspan="2">Verification</td></tr>
<tr><td>Creation of result</td><td colspan="2">(Pure) Data Collection</td><td colspan="2">Data processing</td><td colspan="2">Knowledge generation</td></tr>
<tr><td>Use of result</td><td colspan="2">Publically available</td><td colspan="2">Collective use</td><td colspan="2">Personal use</td></tr>
<tr><td rowspan="5">Campaign</td><td>Duration</td><td colspan="2">Fixed</td><td colspan="2">Implicit limited</td><td colspan="2">Unlimited</td></tr>
<tr><td>Location</td><td colspan="3">Limited</td><td colspan="3">Unlimited</td></tr>
<tr><td>Type of group</td><td colspan="2">Single person</td><td colspan="2">Closed</td><td colspan="2">Open</td></tr>
<tr><td>Monitoring</td><td colspan="3">Periodic</td><td colspan="3">Continuous</td></tr>
<tr><td>Application field</td><td colspan="2">Environment</td><td colspan="2">People</td><td colspan="2">Hybrid</td></tr>
</table>

Specific mobile crowdsourcing-based applications differ with regard to a number of various dimensions. This section mainly focuses on differences of higher level systems that enable general propositions about the design aspects. Due to significant differences between application domains, their specific methodological aspects are not covered in this paper.

4.1 Mobile devices (type, sensors)

Crowdsourcing campaigns usually use mobile devices to collect user data („Bring your own device!“ philosophy). These are typically widely available Internet-capable and GPS-enabled devices such as smartphones or tablets. Cases in which task-specific hardware is needed to gather user data are rare. Mobile crowdsourcing applications differ with regard to the sensors used for data collection. Users typically enter all data manually (free text, forms etc.). But also sensors installed on mobile devices are often used to collect data. There are several sensor technologies for data collection: position sensors (GPS, magnetic field/compass, proximity sensors etc.), motion sensors (accelerometer, gyroscope etc.), and ambient sensors (light, air pressure, temperature, and humidity sensors). Furthermore, data can be collected with cameras, microphones and by using video, audio and image recording. Prospectively, further sensors will be fitted as standard in mobile devices. In special cases, additional sensors are connected to the device, for example to monitor vital functions of the human body (pulse, skin resistance etc.) or to measure air pollution (pollen, exhaust fumes etc.). It is often possible to find combinations of manual entered and collected sensor data (e.g. rare picture and manually-entered textual description of the place where it was found).

4.2 Data (collection, location, transmission, anonymization, processing)

Data capturing can be done manually, automatically, or context-aware [21]: In the manual mode, the participants are personally involved and trigger the collection of data themselves when they detect relevant events (e.g., a severe weather situation). In the automatic mode, the participants are not directly involved and the data is collected automatically by the embedded sensors of the mobile devices (e.g., noise and air pollution). In the context-aware mode, the embedded sensors monitor their environment and activate the manual or automatic capturing function when previously set thresholds are exceeded. Data captured are often location-based [36]. The location coordinates, mostly simple point coordinates, are determined by an embedded location sensor (GPS etc.) that, if necessary, can be validated. Some applications are also capable of capturing line geometries (e.g. distance walked**),** surface geometries or polygons (e.g. disaster area). Data captured have to be transmitted from mobile devices to a central collection server. Data can be transmitted in real time (e.g., via mobile phone networks) or with a delay (e.g., when WLAN is available). Usually, data that are used for communication purposes (SMS, mobile phone network) are transmitted in real time. In offline situations, delayed transmission may be essential to meet security requirements. In such a case, data transmission would start manually (or, if necessary, automatically) when a connection to the Internet is available (e.g., via WLAN) or other means of communication are in use (LAN, flash drive etc.). Data collected are often transferred in anonymous form. However, there are applications that require user and data authentication (e.g., medical applications for monitoring vital functions of the human body). Some applications produce raw unprocessed data (e.g. noise measurement) or forward data to downstream external systems (e.g., weather or flood forecast). These applications often carry out complex data processing and data analysis operations (e.g., aggregation, recognition of duplicates, visualization), and provide users with post-processed data.

4.3 Participation (admission, selection, criteria, registration, evaluation)

The admission of a participant for inclusion into a data collection campaign can be based on the participant's own initiative or on a request by the campaign organizer to a certain person. The latter can be a binding order or a non-binding enquiry. The required a priori knowledge of the user can be low/none (e.g., automatic sensing of traffic), medium (e.g., capturing of damages in urban areas) or high (e.g., observation of natural processes, air pollution reports, wild animal classification). The selection of the people to be requested can be role-based, person-based or location-based or can be carried out by combination of these different factors. Thereby, the participants of the campaign can be fixed at the beginning of the campaign or participants can also be included during the course of the campaign (or selected on demand). The participants can participate anonymously or can be registered users, i.e. known to the organizer. Furthermore, known participants can be in a formal relation to the organizer, i.e. employee, member, partner, customer etc. of an organization. Often the participants are members of the general public ("motivated crowd"). The assessment of the contributions of the participants can be carried out automatically by the system, by other participants or by the organizer (or other people in charge).

4.4 Involvement (degree, task type, creation of result, use of result)

The degree of participation can have a significant influence on the willingness of the participants to contribute to a campaign – with their time and their own device. It can be distinguished between active and passive participation: In case of an active participation the participant is actively involved in the process. The task of the participant then is not limited to the data capturing, but can contain more tasks ranging from development and formulation of goals to the generation of information and reports based on the data collected during the campaign. In case of a passive participation the participant does not have to become active, because the data capturing application runs on his mobile device in the background and collects and transmits data automatically. The data collection tasks can be initial or non-recurring data collections of new data. They also can concern updates, changes or verifications of existing data. The tasks of the participants are not limited to (pure) data collection. They can involve other tasks of the knowledge and result generation process, like further processing of the collected (raw) data (e.g., model calculation and simulation) and evaluation of the results (reports etc.). With regard to the use of the results public, collective or personal projects can be distinguished. In case of a public project, the results are available to the general public, e.g., collection of damages and shortcomings in a city. In case of a collective project a group of people with a shared interest selects data with regard to a common goal, e.g., proof of pollution caused by a factory. In case of a personal project the participants are focused on self-awareness and self-improvement. They exclusively use the results for their personal purposes, e.g. monitoring and optimization of eating habits, sleep, and sport performance. Possibly these data are shared with friends (e.g., via social networks) or (in anonymized form) made available publically, to learn one from each other.

4.5 Campaign (duration, location, group type, monitoring, application field)

Temporally, the data collection campaign can be limited (fixed duration) or unlimited (open-end). The temporal frame of a campaign may also be implicitly defined, e.g. a collection is carried out until a certain amount of data has been gathered. Locally, especially environment-centric campaigns take place within a spatially delimited area, region or place, e.g. capturing of environmental damages or road conditions. In contrast, people-centric applications often are not limited to a certain, spatially delimited area or place. Different types of tasks may require different numbers of participants and forms of groups. Sometimes, only a single person is necessary to accomplish a certain data collection task. In other cases, a closed group with specific knowledge for a problem solution is required. And in other cases, an open group is formed, because the general public is invited to participate in the data collection task [37]. The monitoring of a campaign (activity of the participants, quantity and quality of the collected data etc.) can be carried out continuously or periodically. In the former case, campaign management is a permanent task, like it is required for large campaigns. The latter requires monitoring of the campaign only ad-hoc, on-demand or at certain moments in time. In general, a mobile crowdsourcing campaign and application can be assigned to one of the aforementioned application fields in the area of people-centric and environment-centric applications. But some applications can be assigned to several application fields, e.g. personal health applications (like monitoring of asthma patients), that also consider environmental conditions (e.g. air pollution).

5 Architecture of Mobile Crowdsourcing Systems

From the analysis of existing mobile crowdsourcing applications and architectures the following architecture of a general crowdsourcing solution (see Figure 2) is derived. This architecture provides a more detailed view on the common functional components (and their interfaces) of mobile crowdsourcing applications.

Typical roles within this architecture are: **a) Campaign organizer:** initiates and monitors the targeted data collection effort (crowdsourcing campaign), including the definition of the campaign as well as recruitment, control and coordination of well-suited participants. **b) Participant:** contributes to the geo-crowdsourcing campaign by (voluntary) capturing and sharing geospatial data using their own mobile device. **c)**

End user: accesses and processes the data captured by the participants according to their interests and needs, e.g. integrating, analyzing and remixing this data.

The general system architecture system is divided into two independent runtime systems: **1) Backend system** (server) with the main components *Campaign Management* (campaign planning, participants management, recruiting, tasking, and campaign monitoring) and *Data Management* (pre-processing, storage, processing and provisioning), and **2) Mobile device** (mobile client) with the data capturing and data transfer components.

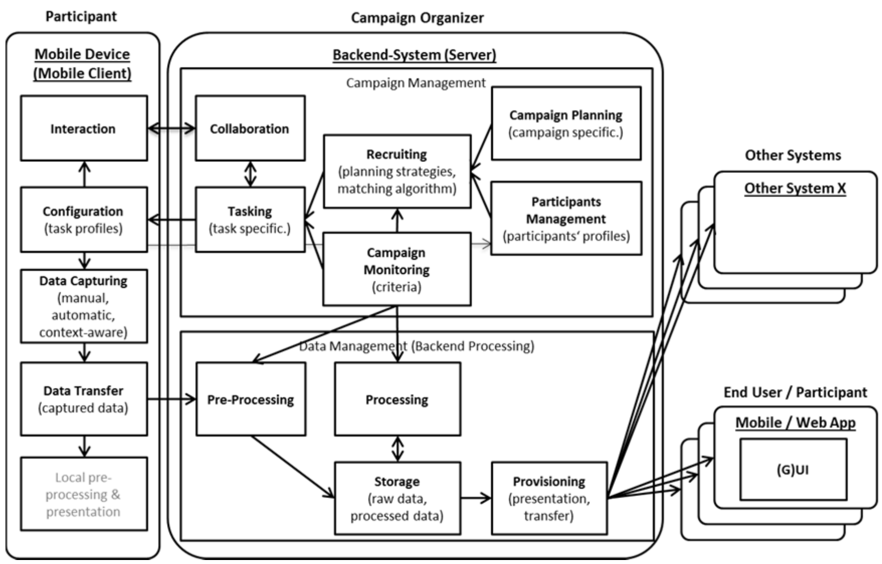

Figure 2. Architectural components and roles of mobile crowdsourcing applications

Typical functional tasks of these architectural components are:

5.1 Campaign Management

Campaign planning: For each data-collection campaign the relevant parameters must be defined in advance. These campaign specifications may involve a number of factors including overall budget, kind of data to be collected (what), geographic coverage (where), timespan, interval or frequency of data collection (when), participants' device capabilities, demographic diversity, and social network affiliation (who).

Participant management: For recruitment, tasking and monitoring of participants certain information about the participants are necessary, that are managed in this component in a participant profile. For recruitment, it is necessary to determine the expertise or skill level of a participant [38]. Relevant characteristics could be participants' device capabilities, reputations as data collectors, motivational structure, skills and availability in terms of geographic and temporal coverage. For authentication registration information is necessary. For tasking, participants' movement pattern could be helpful. The monitoring of participants considers acceptance and rejection decisions of the user's contributions [39]. Additionally, ranking scores can be provided, that present the skill level, the reputation or the quality of the participant [40].

Recruiting: The recruitment component takes campaign specifications and participants' profiles as input and recommends participants for involvement in data collections. For finding a fit between diverse participants (profiles) and the crowdsourcing campaign (specification) certain planning strategies and matching algorithms can be used. In general, participants for campaigns must meet minimum requirements. Advanced approaches identify which subset of individuals is best suited (according to the campaign specification, e.g. budget, maximal coverage etc.).

Tasking: This component distributes the assigned data-collection task to the participants' mobile devices. Tasking can be proactive (the data-collection task is pushed to the participant' device) or passive (user chooses the data collection task from a set of tasks provides by the organizer). The task specifies the data capture modalities based on the campaigns requirements including criteria when, where, what data to capture (e.g., take a 2 megapixel picture at location x in time frame y twice a day).

Campaign monitoring: This component monitors the progress of the campaign according to predefined criteria in the campaign definition such as quantity and quality of the collected (raw and processed) data. The check of the criteria can occur continuously or periodically. If necessary (e.g., a criterion is below a threshold) campaign organizers may be alerted and measures are taken like recruiting additional participants (e.g., if the coverage of a specific area has to be improved or not enough data have been collected) and create additional tasks (e.g., to verify outlier data or to get more data on detected hot spots). In automatic data capturing mode, the tasking action can be completely automated such as changing sampling rates (adaptive sampling) or turning sensors on/off (actuation). In the manual mode, a trigger message can be sent asking participant to collect data on certain location or change route.

Collaboration: This component facilitates collaboration and coordination. Participants need to interact to solve a task collaboratively. Additionally, interaction between participant and organizer may be needed, e.g. to provide feedback to the results of the participant (from organizer to participant), and to ask for more details regarding the task specification of the organizer (from participant to organizer). This component can also be used to assign tasks newly.

5.2 Mobile Device

Configuration: This component manages the assigned tasks on the mobile device of the participant. For each task a profile with the task specification is maintained. The participant can be informed about tasks to fulfill, navigated to locations, where data should be captured, or the sensors of the participant's mobile device can be controlled for data capture manually and automatically. In the configuration component, participants can select different data transfer strategies. Participants also can manage parts of their participant profile (role, place, time etc.) to update and specify the kind of tasks they are willing to fulfill (light grey line to “participant management” in fig. 2).

Interaction: This component provides functionalities to facilitate interaction among participants included in the task profile (in case of a collaborative task) and between participant and organizer. These are functionalities to communicate and share information, manage personal identities, maintain relationships, re-assign tasks among the participants or collaboratively document knowledge using the servers' collaboration component.

Data capturing: This component collects different kinds of data requested in the campaign definition and task profile respectively using the mobile device's data-input mechanisms and built-in or plugged-in sensors - prevalently camera (picture, video), microphone (audio), clock (time), GPS (location), gyroscope, accelerometer etc. Capturing can be done manually, automatically, or context-aware (see 4.2).

Data transfer: This component ensures the transmission of the data collected by the capturing component to the backend system (server). The data can be transferred immediately or delayed (see 4.2). This component also ensures short-time storage of the data to be transmitted to the backend system (server).

5.3 Data Management

Pre-processing: This component analyzes the transmitted data and prepares them for storage. Usually, several application-specific processing steps are necessary. But in many cases it may also be very simple. Examples of those steps are: a) Data extraction and transformation of the data into the internal data structure, for instance by applying audio analysis (speech recognition to extract words, sound classification to match sounds, for instance for detecting birds voices etc.) or image analysis (e.g., determining the water level with a picture of a gauge board etc. or optical character recognition to scan text, object recognition to find objects). b) Data assessment and analysis, for instance data cleaning, quality assessment with data filtering (eliminating duplicates etc.), data fusion and aggregation of multiple or similar data.

Storage: This component ensures the long-term storage of the (raw and processed) data. The data is stored in relational databases or databases specially adapted the management of spatial data or sensor data. In some cases a history and a comparison of current data with historical trends are necessary, so that a robust long-term storage is a central requirement.

Processing: This component processes the stored raw data to extract features of interest und get insights about the observed phenomenon. This step is typically application specific and potentially involves an immense number of (spatial) data processing methods ranging from numeric modeling to descriptive statistics to image processing and sophisticated machine learning algorithms.

Provisioning: This component prepares the data for presentation or for transfer to other systems. In case of a standalone system this component provides presentation and visualization services to present the results obtained by the processing component to the end user. The results can be visualizations of the raw data or the processed data and are usually presented through web-portals or on mobile devices (often on maps). In case of external systems interested in using the data, this component provides services to access the data including necessary transformations.

6 Example Applications

The architecture proposed in section 5 was successfully applied in several projects. Two examples from the field of disaster management are described in this section.

6.1 Flood Risk App "Hochwasserrisiko"

Floods belong to the natural disasters which regularly take human life and cause high damages to property [41]. In order to take precautionary measures for protection, the authorities responsible for and the citizens endangered by flood need up-to-date information on the water levels and the actual state of risk [42]. In the first example, a community of volunteer water level and gauge observers transmits via a smartphone app the geo-referenced measurements and a photo of the scale of non-automatic gauges to the authority

With the aid of the app "Hochwasserrisiko" [43] the current water and gauge levels at rivers can be collected and entered by interested citizens themselves with a photo and metadata (1 & 2 in Fig.3). This information can be displayed on a map or in augmented reality in the mobile devices of the users (3 & 4 in Fig. 3). In addition, this information is further processed by the flood authority's warning center for more precise flood prognosis and warning. These flood warnings again can be displayed on the mobile devices of the users (red icon in 3 & 4 in Fig. 3). In addition, with this data flood risk maps [44] can be processed, which inform about the current risk potential of floods (5 & 6 in Fig. 3). They can be viewed on a map (bird's eye view) or through the smartphone's live camera (augmented reality view). Using augmented reality, the virtual water level of a flood scenario can be visualized in reality. This allows a realistic display and improves the perception and analysis of the scenario's hazard.

Figure 3. Screens of the flood risk app "Hochwasserrisiko"

This scenario works with a fixed set of a priori selected citizens, which receive an order with detailed instructions, where to collect water and gauge levels. The users can see the places where the information is needed on a map (see 1 in Fig. 3) and select and accept the places (tasks). The users then move to the places selected and collect the information manually by entering in a form and taking some pictures (see 2 in Fig. 3). This data is then transferred to a central server of the flood authority's warning center, where this information is controlled by an operator who compares the photo with the measurements, thus ensuring the same quality of measurements as with automatic gauges. In this case the data filtering and fusion is performed by an operator. This is necessary because the crowdsourcing data is used as an input for further model processing. This is only applicable when the data characteristics and the data quality are adequate for being used as a basis for further algorithmic hazard prediction methods, in other words when it can be assured that the mobile crowdsourcing data can substitute the data of a physical sensor. The processed data (flood warnings and flood risk maps) is than available on the mobile devices of the users and the general public (see 3-6 in Fig. 3).

6.2 Disaster management app "Emergency help"

Modern communication technologies such as mobile applications, that inform people about natural disasters occurring in their vicinity, are increasingly gaining in importance. For example, the KATWARN app offers location-based disaster warnings for small areas. KATWARN warnings are issued by the local emergency management agencies such as fire departments [45]. Mobile crowdsourcing offers possibilities for sustained improvement of danger prevention processes.

The app „Emergency help", which is an expansion of the KATWARN app, helps to manage disaster relief volunteers, which are skilled to provide support and assistance, for example trained in first aid. The app informs people in the affected area about an upcoming hazard and how they should behave to stay safe, for example to leave the danger zone. Moreover, it allows preregistered disaster relief volunteers to provide certain information about their skills (first aid, languages etc.), to signal their willingness to give assistance in case of a hazard, and to inform about the completion of assigned assistance tasks. Another user group of the app is people in need of help that are not able to follow all security procedures, e.g. elderly people with mobility impairments who cannot leave the danger zone on their own. The „Emergency help" app offers such people the possibility to provide information on their physical and health problems, and to signal their need for aid in case of a hazard.

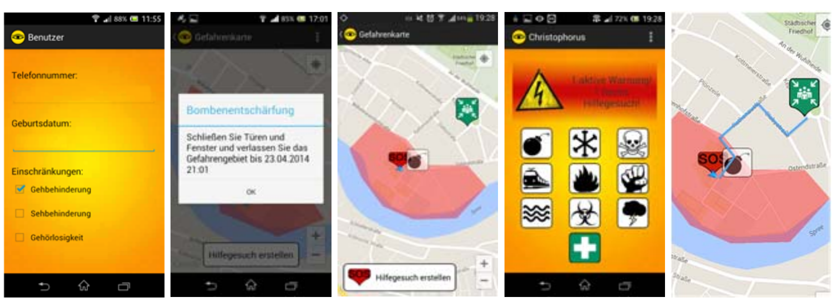

Figure 4. Screens of disaster management app "Emergency help"

After a warning is sent, an additional button is displayed in the app. People in need of help can use this button to request help. The person's disability information and location data are sent to the app's central service. There, based on the request received, the service searches for voluntary helpers in the vicinity. The helper selected receives a notification that his assistance is needed. The helper can now confirm his willingness to assist. Then, he receives the (anonymized) data of the person in need, e.g. his or her location. Based on the exchanged data the app supports the helper, e.g. information sharing with the person in need (name or address if needed), navigation to the location of the person in need or evacuation routes. The app shows the person in need that the request was accepted, and who and how far the helper is. The successful completion of a rescue operation is reported to the central service. , so the responsible emergency management agencies have an overview about all registered people in a case of a crisis.

In this scenario, a dynamically forming group of skilled and unskilled participants is participating in the campaign ad-hoc, voluntary and possibly anonymous, based on the participants own initiative. Thereby, point coordinates (location of the user, which is provided by the mobile standard device's GPS localization function) are automatically captured in real-time, anonymized, and transmitted with other manually captured, structured data (profile, incl. role of user). These raw data are not further processed on the server's side, but only used for the coordination functionality provided by the system. Temporally, the campaign is implicitly limited to the duration of the dangerous situation. Locally, the campaign is spatially delimited to the affected area.

In this scenario, the campaign management subsystem plays a crucial role. The configuration component is used to specify the role of the user (helper or in need) and updates the participant profile in the participant management component. The participant management and the recruiting component are important for managing, selecting and matching the disaster relief volunteers. The tasking component then provides the necessary information for coordination among the people involved, which interact using the interaction and collaboration component. The data capturing component only collects data about location of user and completeness of assistance tasks. Campaign monitoring provides information about the progress and monitors people still in need of help.

7 Summary and Outlook

In this paper we investigated different dimensions of and typical functionalities of mobile crowdsourcing systems. The main goal of this paper was to gain a better understanding about design aspects to be considered in the development of such systems. Therefore, the paper presented a classification scheme and a general architecture with the typical roles, components and functionalities of mobile crowdsourcing systems. The aim was to generalize and study mobile crowdsourcing applications, and to facilitate the understanding of current systems and the design of new ones.

On the surface these systems appear deceptively simple: participants are asked to collect data, the data is collected and then analyzed for a variety of purposes. But the current diversity of mobile crowdsourcing applications, which is represented within the classification scheme, shows the difficulties and the complexity, when designing and implementing such an application. This requires not only profound knowledge how mobile crowdsourcing works and when it can be applied [46], but also how it is technically designed and implemented. Therefore, the general architecture may be used as a blueprint and to guide decisions during the development process.

However, the design and deployment of successful systems require meeting simultaneously a diverse set of technical and people-centric challenges. Even though (mobile) crowdsourcing has been a vital research area over the last several years, there are still several research questions in the field, e.g. data quality (potential for accidental submission of bad data or malicious submissions) [47], [48], privacy concerns of the users (risk of loss or theft of the device with personal data, profiling, central data archives) [30], and the motivation issues that participants may have within the system [49]. Other issues are related to the scalability of the technical infrastructure and the campaign organization [50].

8 References

1. Greengard, S.: Following the crowd. Communications of the ACM. 54, pp. 20–22 (2011)
2. Mea, V.D., Maddalena, E., Mizzaro, S.: Crowdsourcing to Mobile Users - A Study of the Role of Platforms and Tasks. In: Cheng, R., Sarma, A.D., Maniu, S., Senellart, P. (eds.): Proceedings of First VLDB workshop on databases and crowdsouring (DBCrowd2013), Volume 1025, CEUR, pp. 14-19 (2013)
3. Chatzimilioudis, G.; Konstantinidis, A.; Laoudias, C.; Zeinalipour-Yazti, D.: Crowdsourcing with Smartphones. In: Internet Computing, IEEE, vol.16, no.5, pp. 36-44 (2012)
4. Chon, Y., Lane, N. D., Kim, Y., Zhao, F., Cha, H.: A Large-scale Study of Mobile Crowdsourcing with Smartphones for Urban Sensing Applications. UbiComp'13 (2013)
5. Abecker, A., Braun, S., Kazakos, W., Zacharias, V.: Participatory Sensing for Nature Conservation and Environment Protection. In: 26th International Conference on Informatics for Environmental Protection (EnviroInfo2012), Shaker, Aachen, pp. 393-401 (2012)
6. Goodchild, M.F.: Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. International Journal of Spatial Data Infrastructures Research 2, pp. 24-32 (2007)
7. Sieber, R.: Public participation geographic information systems: A literature review and framework. Annals of the American Association of Geography, 96(3), pp. 491–507 (2006)
8. Burke, J., Estrin, D., Hansen, M., Parker, A. Ramanathan, N. Reddy, S., Srivastava, B.: Participatory sensing. In: Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications (Oct. 2006), pp. 117–134 (2006)
9. Zhao, Y., Zhu, Q.: Evaluation on crowdsourcing research: Current status and future direction. Information Systems Frontiers. Pp. 1-18 (2012)
10. Brabham, D.C.: Crowdsourcing as a Model for Problem Solving - An Introduction and Cases. Convergence 14(1), pp. 75-90 (2008)
11. Estellés-Arolas, E., González-Ladrón-de-Guevara, F.: Towards an integrated crowdsourcing definition. Journal of Information Science 38(2), pp. 189–200 (2012)
12. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R., Schader, M.: Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. In: 17h Americas Conference on Information Systems. pp. 1–11 (2011)
13. Rouse, A.C.: A preliminary taxonomy of crowdsourcing. In: ACIS 2010, paper 76 (2010)
14. Schenk, E., Guittard, C.: Towards a characterization of crowdsourcing practices. Journal of Innovation Economics 1(7), pp. 93-107 (2010)

15. Erickson, T.: Some Thoughts on a Framework for Crowdsourcing. In: CHI 2011 Workshop on Crowdsourcing and Human Computation. pp. 1-4 (2011)
16. Yuen, M.-C., King, I., Leung, K.-S.: A Survey of Crowdsourcing Systems. Privacy, security, risk and trust. In: 3rd Int. Conf. on Social Computing, IEEE, pp. 766–773 (2011)
17. Doan, A., Ramakrishnan, R., Halevy, A.Y.: Crowdsourcing systems on the World-Wide Web. Communications of the ACM. 54, pp. 86–96 (2011)
18. Kazman, R., Chen, H.-M.: The metropolis model a new logic for development of crowdsourced systems. Communications of the ACM. 52, pp. 76–84 (2009)
19. Malone, T.W., Laubacher, R., Dellarocas, C.: The collective intelligence genome. Engineering Management Review, IEEE. 38, pp. 38–52 (2010)
20. Hetmank, L.: Components and Functions of Crowdsourcing Systems – A Systematic Literature Review. In: 11th Int. Conf. on Wirtschaftsinformatik, Leipzig, pp. 55-69 (2013)
21. Estrin, D.: Participatory Sensing: Applications and Architecture. In: 8th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys), pp. 3-4 (2010)
22. Tilak, S.: Real-World Deployments of Participatory Sensing Applications: Current Trends and Future Directions. ISRN Sensor Networks, vol. 2013, Article ID 583165 (2013)
23. Khorashadi, B., Das, S.M., Gupta, R.: Flexible architecture for location based crowdsourcing of contextual data. Patent US 8472980 B2 (2011)
24. Reddy, S., Estrin, D., Srivastava, M.: Recruitment Framework for Participatory Sensing Data Collection. In: Int. Conference on Pervasive Computing (Pervasive’2010), (2010)
25. Yan, T., Marzilli, M., Holmes, R., Ganesan, D., Corner, M.: mCrowd - A Platform for Mobile Crowdsourcing. In: SenSys’09, Berkeley, USA, pp. 347-348 (2009)
26. Lasnia, D., Bröring, A., Jirka, S., Remke, A.: Crowdsourcing Sensor Tasks to a Socio-Geographic Network. In: 13th AGILE Conf. on Geographic Information Science (2010)
27. Das, T., Mohan, P., Padmanabhan, V., Ramjee, R., Sharma, A.: PRISM: Platform for Remote Sensing Using Smartphones. In: 8th Int. ACM Conference on Mobile Systems, Applications, and Services (MobiSys’10), San Francisco, California, USA (2010)
28. Luqman, F., Griss, M.: Overseer: A Mobile Context-Aware Collaboration and Task Management System for Disaster Response. In: 8th Int. Conf. on Creating, Connecting and Collaborating through Computing (2010)
29. Christin, D., Reinhardt, A., Kanhere, S., Hollick, M.: A survey on privacy in mobile participatory sensing applications. J. of Systems and Software 84 (11), pp. 1928-1946 (2011)
30. Cristofaro, E. de, Soriente, C.: PEPSI - Privacy-Enhanced Participatory Sensing Infrastructure. ACM Conf. on Wireless Network Security (WiSec'11), pp. 23-28 (2011)
31. Ra, M., Liu, B., La Porta, T., Govindan, R.: Medusa: A Programming Framework for Crowd-Sensing Applications. In: 10th International Conference on Mobile Systems, Applications, and Services (MobiSys'12), pp. 337-350 (2012)
32. Hupfer, S., Muller, M., Levy, S., Gruen, G., Sempere, A., Ross, S., Priedhorsky, R.: MoCoMapps: mobile collaborative map-based applications. In: Conf. on Computer Supported Cooperative Work (CSCW '12). ACM, New York, NY, USA, pp. 43-44 (2012)
33. Cornelius, C., Kapadia, A., Kotz, D., Peebles, D., Shin, M., Triandopoulos. N.: Anonysense: Privacyaware people-centric sensing. In: ACM MOBISYS, pp. 211-224 (2008)
34. Kitchenham, B.: Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report, Keele University, Keel, UK (2007)
35. Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology. 51, Elsevier, pp. 7–15 (2009)
36. Gonzalez, A.L., Izidoro, D.; Willrich, R., Santos, C.A.S.: OurMap: Representing Crowdsourced Annotations on Geospatial Coordinates as Linked Open Data. In: Antunes,

P., Gerosa, M.A., Sylvester, A., Vassileva, J., Vreede, G.-J. (eds.): Collaboration and Technology - CRIWG 2013, LNCS, vol. 8224, pp. 77-93 (2013)
37. Fraternali, P., Castelletti, A., Soncini-Sessa, R., Ruiz, C.V., Rizzoli, A.E.: Putting humans in the loop: Social computing for Water Resources Management. Environmental Modelling; Software. 37, pp. 68–77 (2012)
38. Corney, J.R., Torres-Sánchez, C., Jagadeesan, A.P., Yan, X.T., Regli, W.C., Medellin, H.: Putting the crowd to work in a knowledge-based factory. Advanced Engineering Informatics. 24, pp. 243–250 (2010)
39. Mashhadi, A.J., Capra, L.: Quality control for real-time ubiquitous crowdsourcing. In: 2nd International Workshop on Ubiquitous Crowdsouring. ACM, New York, pp. 5-8 (2011)
40. Schall, D.: Expertise ranking using activity and contextual link measures. Data & Knowledge Engineering. 71(1), pp. 92–113 (2012)
41. Blöschl, G., Montanari, A.: Climate change impacts - throwing the dice? Hydrological Processes 24(3), pp. 374–381 (2010)
42. Hornemann, C., Rechenberg, J.: Was Sie über den vorsorgenden Hochwasserschutz wissen sollten. Umweltbundesamt, Dessau (2006)
43. Fuchs-Kittowski, F. Simroth, S., Himberger, S., Fischer, F.: A content platform for smartphone-based mobile augmented reality. In: Arndt, H.-K.; Knetsch, G.; Pillmann, W. (Hrsg.): 26th International Conference on Informatics for Environmental Protection (EnviroInfo2012) – Part 1: Core application areas. Shaker, Aachen, pp. 403-412 (2012)
44. Europäische Gemeinschaft: Richtlinie 2007/60/EG des Europäischen Parlaments und des Rates vom 23. Oktober 2007 über die Bewertung und das Management von Hochwasserrisiken. ABl. L 288 vom 06.11.2007 (2007)
45. Meissen, U., Faust, F., Fuchs-Kittowski, F.: WIND - A meteorological early warning system and its extensions towards mobile devices. In: Page, B. et al. (Hrsg.): 27th Int. Conf. on Environmental Informatics (EnviroInfo 2013), Shaker, Aachen, pp. 612-619 (2013)
46. Thuan, N.H., Antunes, P., Johnstone, D.: Factors Influencing the Decision to Crowdsource. In: Antunes, P., Gerosa, M.A., Sylvester, A., Vassileva, J., Vreede, G.-J. (eds.): Collaboration and Technology CRIWG2013, LNCS, vol. 8224, pp. 110-125 (2013)
47. Jordan, L., Stallins, A., Stokes, S., Johnson, E. Gragg, R.: Citizen Mapping and Environmental Justice: Internet Applications for Research and Advocacy. Environmental Justice 4(3), pp. 155–162 (2011)
48. Thogersen, R.: Data Quality in an Output-Agreement Game - A Comparison between Game-Generated Tags and Professional Descriptors. In: Antunes, P., Gerosa, M.A., Sylvester, A., Vassileva, J., Vreede, G.-J. (eds.): Collaboration and Technology CRIWG 2013, LNCS, vol. 8224, pp. 126-142 (2013)
49. Vreede, T. de, Nguyen, C., Vreede, G.-J. de, Boughzala, I., Oh, O., Reiter-Palmon, R.: A Theoretical Model of User Engagement in Crowdsourcing. In: Antunes, P., Gerosa, M.A., Sylvester, A., Vassileva, J., Vreede, G.-J. (eds.): Collaboration and Technology CRIWG 2013, LNCS, vol. 8224, pp. 94-109 (2013)
50. Liu, C.H., Hui, P., Branch, J.W., Bisdikian, C., Yang, B.: Efficient network management for context-aware participatory sensing. In: 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON) pp. 116-124 (2011)