Chapter 4
Human Sensors

**Vassilis Kostakos, Jakob Rogstadius, Denzil Ferreira,
Simo Hosio, and Jorge Goncalves**

4.1 Introduction

This chapter provides an overview of how humans can themselves act as sensors—
human sensing—for data collection by considering a variety of scenarios. We
start by providing a survey of literature, followed by proposed guidelines, and
lastly discussing case studies which exemplify using humans for sensing and data
collection. We present human sensing in three technical domains:

1. *Human sensors online*, where we describe how online crowd markets are
   enabling the aggregation of online users into working crowds, and discuss
   important motivation techniques and strategies for this topic.
2. *Online social media mining* on a large scale, where we exemplify how users'
   posting of opinions and content in online social media is enabling us to develop
   platforms that analyse and respond to this data in realtime. Crisis response
   systems are a very popular type of system in this category, and here we present
   an overview of many of these systems, along with a case study that focuses on
   one of these systems called *CrisisTracker*.
3. *Offline human sensors,* i.e. urban and in-situ systems that collect data from
   pedestrians. Here we provide an overview of crowdsourcing beyond the desktop,
   and of systems that are designed to collect opinions from pedestrians in an

V. Kostakos (✉) • D. Ferreira • S. Hosio • J. Goncalves
Center for Ubiquitous Computing, University of Oulu, Oulu, Finland
e-mail: vassilis@ee.oulu.fi; denzil.ferreira@ee.oulu.fi; simo.hosio@ee.oulu.fi;
jorge.goncalves@ee.oulu.fi

J. Rogstadius
Madeira Institute of Interactive Technologies, University of Madeira, Funchal, Portugal
e-mail: jakob.rogstadius@m-iti.org

urban context. We present a case study on a set of systems that use public displays to collect feedback from citizens, and provide strategies and guidelines for conducting this kind of work.

4.2 Human Sensors Online

One way to rely on humans as sensors is to collect data from them directly. Increasingly, large numbers of online users are aggregating in online markets, like Amazon's Mechanical Turk or Crowdflower, to making themselves accessible to anyone who is interested in rewarding their time for completing some task online. Online crowd markets are generic enough, and a variety of tasks can be completed by workers in such markets, ranging from answers to surveys to writing restaurant reviews, movie reviews, annotating photographs, transcribing audio, and any other task which computers cannot reliably do at the moment—at least without training data obtained from humans first. A great example of using these online markets for sensing is *Zensors*, which enables creation of arbitrary sensors for any visually observable property (Laput et al. 2015). In practice Zensors sends images for the crowds to process and label according to clear instructions on what to look for. Using the markets, Zensors is able to produce near-instant sensor readings about the properties, and once enough data has been collected the results can be handed off to a machine learning classifier for automated sensing in future cases. These markets can indeed be an important source of collecting data from humans, but the fact that they are structured as a market (as opposed to, say Facebook) has important implications for motivating and attracting people to certain tasks.

4.2.1 Crowdsourcing Markets and Mechanical Turk

A number of crowdsourcing markets exists, with the one of most studied being Amazon's Mechanical Turk, a general marketplace for crowdsourcing where requesters can create Human Intelligence Tasks (HITs) to be completed by workers. Typical tasks include labelling objects in an image, transcribing audio, or judging the relevance of a search result, with each task normally paying a few cents (USD). Work such as image labelling can be set up in the form of HIT groups, where the task remains identical but the input data on which the work is carried out varies. Mechanical Turk provides a standardized workflow within such groups where workers are continuously offered new tasks of the same type after they complete a task within the group. Mechanical Turk also allows duplicating a HIT into multiple identical assignments, each of which must be completed by a different worker, to facilitate for instance voting or averaging schemes where multiple workers carry out the same task and the answers are aggregated.

4.2.2 Motivating Workers in Crowdsourcing Markets

It is important to provide an overview of why people take part as workers in crowdsourcing markets, and what does the theory suggest about their performance in completing tasks. A traditional "rational" economic approach to eliciting higher quality work is to increase extrinsic motivation, i.e., an employer can increase how much they pay for the completion of a task (Gibbons 1997). Some evidence from traditional labor markets supports this view: Lazear (2000) found workers to be more productive when they switched from being paid by time to being paid by piece; Hubbard and Palia (1995) found correlations between executive pay and firm performance when markets were allowed to self-regulate.

However, there is also evidence that in certain situations financial incentives may not help, or may even hurt. Such extrinsic motivations may clash with intrinsic motivations such as a workers' desire to perform the task for its own sake. This is particularly important in the context of online crowdsourcing where the "employer" does not control the working environment of workers.

For example, a classic experiment by Deci (1975) found a "crowding out" effect of external motivation: students paid to play with a puzzle later played with it less and reported less interest than those who were not paid to do so. In the workplace, performance-based rewards can be "alienating" and "dehumanizing" (Etzioni 1971). If the reward is not substantial, then performance is likely to be worse than when no reward is offered at all; insufficient monetary rewards can act as a small extrinsic motivation that tends to override the possibly larger effect of the task's likely intrinsic motivation (Gneezy and Rustichini 2000). Given that crowdsourcing markets such as Mechanical Turk tend to pay very little money and involve relatively low wages (Paolacci et al. 2010), external motivations such as increased pay may have less effect than requesters may desire. Indeed, research examining the link between financial incentives and performance in Mechanical Turk has generally found a lack of increased quality in worker output (Mason and Watts 2009). The relationship between price and quality has also had conflicting results in other crowdsourcing applications such as answer markets (Harper et al. 2008). Although paying more can get work done faster, it has not been shown to get work done better.

Another approach to getting work done better could be increasing the intrinsic motivation of the task. Under this view, if workers find the task more engaging, interesting, or worth doing in its own right, they may produce higher quality results. Unfortunately, evidence so far regarding this hypothesis has been conflicting. For example, work by Chandler and Kapelner (2013) reported that while crowdsourcing tasks framed in a meaningful context motivate individuals to do more, they are no more accurate. On the other hand, work by (Rogstadius et al. 2011a) suggests that intrinsic motivation has a significant positive effect on workers' accuracy, but not productivity.

These contradictory results and a number of other issues that suggest the question of motivating crowd workers has not yet been definitively settled. First, prior studies

have methodological problems with self-selection, since workers may see equivalent tasks with different base payment or bonuses being posted either in parallel or serially. Second, very few studies besides have looked at the interaction between intrinsic and extrinsic motivations; Mason and Watts (2009) vary financial reward (extrinsic), while Chandler and Kapelner (2013) vary meaningfulness of context (intrinsic) in a fixed diminishing financial reward structure. Finally, the task used in Chandler and Kapelner (2013) resulted in very high performance levels, suggesting a possible ceiling effect on the influence of intrinsic motivation.

4.2.3   Running Experiments on Mechanical Turk

Using Mechanical Turk has posed a problem for experimental studies, since it lacks support for random participant assignment, leading to issues even with between subjects control. This is especially problematic for studies of motivation, as self-selection is an inherent aspect of a task market. This means that results in different conditions could be due to attracting different kinds of people rather than differences in the conditions themselves.

For example, given two tasks of which one pays more and one pays less, making both of them available on the site at the same time would bias the results due to the contrast effect. This contrast effect would be problematic even for non-simultaneous posting if workers saw one task at one price and then the same task at another price at a later time. If tasks were put up at different times, then different workers might be attracted (e.g., Indian workers work at different times than Americans; some days/times get more activity than others, etc.), or more attractive work could be posted by another requester during one of the conditions but not the other.

The other extreme is to host everything on the experiment server, using Mechanical Turk only as a recruitment and fulfilment host. All participants see and accept the same identical task, and are then routed to the different places according to the appropriate condition on the experimenter's side. This fails when studying how workers act naturalistically, as everything is on the host environment. Thus aspects such as the title, description, and most importantly reward cannot be varied by condition, making it impossible to study natural task selection.

For these reasons, an approach proposed by (Rogstadius et al. 2011a) was for participants to fill out a common qualification task with neutral title and description. This qualification task (for example, simply collecting demographic data) is hosted on the researcher's server (rather than Mechanical Turk), and on completion randomly assigns the participant to one of the conditions through a "condition-specific qualification" in the Mechanical Turk system. This qualification enables workers to see and select only tasks in that condition when searching for tasks in the natural MTurk interface. In their study, Rogstadius et al. (2011a) used a Mechanical Turk qualification type with six different possible values corresponding to the different conditions. The key benefit of this approach is that participants still use the Mechanical Turk interface as they naturally do to self-select tasks, which can

have condition-specific titles, descriptions, content, and rewards. While participants can still explicitly search for the tasks in other conditions and see them in some HIT listings, HITs cannot be previewed without having the appropriate qualification. Hosting the task externally would avoid the explicit search problem, but would not address non-preview textual descriptions or the key issue of supporting condition-specific variations in payment.

Another advantage of the qualification-task-approach is that the worker will always retain the qualification granted to them by the experimenter (so they can be kept track of). Thus, for example if an experimenter wanted to make a new experiment available to a subset of their participants they could add the qualification for it to the appropriate participants and the task would automatically become available to the target participants on Mechanical Turk. For more intensive recruitment, once a worker has completed the qualification task and their worker ID is known, they can be emailed directly by the experimenter, even if they did not complete an experiment.

This proposed approach for recruiting participants from a crowdsourcing market lets us retain some of the control of a traditional laboratory setting, the validity of participants searching for work in their natural setting, and the benefits offered by a greater diversity of workers more representative of the online population than undergraduates would be (Horton et al. 2011). The legitimacy of doing both cognitive and social experiments with Mechanical Turk has been supported by multiple studies, e.g. (Heer and Bostock 2010; Paolacci et al. 2010).

4.2.4 Strategies and Guidelines for Crowdsourcing

A number of strategies are proposed by Rogstadius et al. (2011a) on how to conduct experiments using Mechanical Turk, which we summarise here. The importance of adequate payment on a crowdsourcing market like Mechanical Turk is crucial. For example, they report that higher paying tasks attract workers at a higher rate, and that those workers also completed more work once they showed up. This resulted in both higher and more predictable rates of progress. The effect which payment has on progress is simple: higher payment leads to quicker results. In addition to increased payment, their data showed that quicker results can be achieved by simplifying each work item, which in turn increases uptake of workers. Finally, they found that no effect of intrinsic motivation on work progress. However, uptake might be improved by highlighting intrinsic value in task captions and summaries as well.

Emphasizing the importance of the work has also been shown to have a statistically significant and consistent positive effect on quality of answers in the same study. By varying the level of intrinsic motivation they show that this effect is particularly strong at lower payment levels, with differences in accuracy of 12 and 17 % for tasks worth 0 and 3 cents respectively. This difference between conditions was even more conservative than Chandler and Kapelner (2013), who either gave workers a description of purpose or did not. These results have application to

crowdsourcing charity work, suggesting that lower payment levels may produce higher quality results. It is unlikely that workers actually prefer to work for less money, thus this might suggest that intrinsic value has to be kept larger than extrinsic value for the accuracy benefits to appear. Clearly a number of other factors may also affect intrinsic motivation including social identity, goal setting, and feedback (Beenen et al. 2004; Cosley et al. 2005).

4.3 Social Media Mining

In this section we provide an overview of online *social media mining* on a large scale. These are systems that consider how users' posting of opinions and content in online social media can enable us to gain insights into unfolding events. We survey a variety of online systems that collect user contributions, and summarise a few ways in which analysis and mining of such data can be seen as a sensor of human behaviour. Finally, we focus on systems that conduct real-time analyses of such data. Crisis response systems are a very popular type of system in this category, and here we present an overview of many of these systems, along with a case study that focuses on one of these systems called *CrisisTracker*.

4.3.1 End-User Contributions as Sensor Data

The widespread availability of smartphones and high-speed connectivity has enabled a range of systems that collect a variety of different types of user contributions. Some of the most popular websites on the Internet now allow people to upload content: YouTube allows users to upload videos, Flickr hosts photographs, and Facebook allows a variety of media and additionally lets people tag this media with relevant keywords. While obviously the original purpose of this content is different, the freely accessible user generated content can be regarded and processed as sensor data, originating from end-users.

Providing a system that allows users to easily tag objects can result in a valuable repository of knowledge. For example, the Wheelmap system allows users to tag, and also search for, wheel-chair accessible places using ones phone ("Wheelmap" n.d.), and in fact research suggests that doing so influences one's own views on accessibility and disabilities (Goncalves et al. 2013b). Other systems allow users to provide location-based recommendations for restaurants or similar venues. Some examples of this include giving location-aware recommendations for restaurants (Alt et al. 2010), or even providing a real-time news report from the place in which they are (Väätäjä et al. 2011). At the same time, researchers are exploring ways in which mobile phones can enable a new empowering genre of mobile computing usage known as Citizen Science (Paulos et al. 2008). Citizen Science can be used collectively across neighborhoods and communities to enable individuals to become

active participants and stakeholders. More broadly, efforts such as the OpenStreet Maps allows users to annotate publicly available maps by adding shops, streets, and landmarks that are missing from the map. Commercially-driven services such as FourSquare and Google plus allow owners of businesses to add their business to the map services, and annotate it with the various facilities it offers.

In addition, recent work has shown how technology can be used to automatically and passively generate tags. For example, one project showed how smartphones in cars can use their accelerometers collectively to find potholes and problematic road surfaces (Perttunen et al. 2011). The simple premise of this work is that all phones travelling in a car will sense a "bump" when they go over a pothole, and so a combination of GPS and accelerometer data can be used to identify such problematic locations. By deploying this simple technology on taxis, buses, or other transport vehicles that routinely travel in a city, a good portion of the street network can be surveyed relatively inexpensively.

Finally, technology in general can be used to tag our own everyday lives and events, for example using the various sensors on smartphones to tag photographs with contextual information (Qin et al. 2011), and more broadly capturing an increasing aspect of our daily routines (Nguyen et al. 2009). This new abundance of everyday information about and around us opens up several avenues for new applications and research in general. One of the popular means to refine this data into something useful, into higher-level abstractions, is to leverage machine learning.

4.3.2 Machine Learning as a Sensor

Machine learning explores algorithms to learn from and make predictions on many types of data. A typical approach is to process an initial set of training data to learn from, and then predict future events or make data-driven decisions based on the historical data. Recent advances in the analysis of the large datasets amassed online are hinting at the true potential of using these techniques as a sensor of large-scale human behaviour. The range of data collected online is ever increasing, and a number of projects demonstrate how this data can act as a sensor & predictor of human activity.

A frequent domain within which machine learning techniques are applied is Twitter. For instance, researchers have shown how a sentiment analysis of the posts made on Twitter can be used to predict the stock market (Bollen et al. 2011). A reason why this works is because Twitter acts as a repository of sentiments and moods of the society, which have also been shown to affect investors in the stock market. Therefore, sentiment analysis of Twitter feed can be used as sensor to predict stock market activity. Similarly, research has shown how an analysis of Twitter can predict how well movies perform in the box office (Rui et al. 2013). More broadly speaking, due to the ephemeral nature of Twitter communications,

users' moods and tendencies appear to leave a "fingerprint" on Twitter itself, and careful analysis of this data can help in predicting real-life outcomes.

A second major source of predicting real-world outcomes is online searches. Already from 2008, researchers had shown that the search volume of influenza-related queries on Yahoo can help predict the outbreak of influenza (Polgreen et al. 2008). The same finding has been subsequently verified with the search volume of queries by Google (Dugas et al. 2012; Ginsberg et al. 2009), and more broadly research suggests that web searches can predict consumer behaviour in general (Goel et al. 2010). At an even more fundamental level, recent work showed that Google search volume correlates with the volume of pedestrian activity (Kostakos et al. 2013), meaning that spikes in the Google searches relating to names of locations, places, or organisations, correlate with spikes in the number of people who physically visit such locations.

The increasing availability of large datasets online suggests that more and more of the events happening in the real world can be predicted, or possibly understood, through a careful analysis of the online traces that our societies are generating. As we describe next, achieving a real-time analysis capability of this data can provide great benefits.

4.3.3   Realtime Mining of Social Media

Social media are used in the emergency response cycle to detect potential hazards, educate citizens, gain situation awareness, engage and mobilize local and government organizations and to engage volunteers and citizens to rebuild the environment. Users of social media at disaster time include victims, volunteers, and relief agencies. Existing systems can be loosely grouped into disaster management ("Sahana Foundation" n.d.; "VirtualAgility WorkCenter" n.d.), crowd-enabled reporting (Rogstadius et al. 2013a; "Ushahidi" n.d.) and automated information extraction (Abel et al. 2012; Cameron et al. 2012; Steinberger et al. 2013).

Sahana ("Sahana Foundation" n.d.) and VirtualAgility OPS Center (VOC) ("VirtualAgility WorkCenter" n.d.) support the emergency disaster management process with information and inventory management and collaboration support for response organizations (emergency teams, security, social workers, etc.) Such systems often integrate raw social media feeds, but typically lack capabilities for distilling and handling useful reports, and avoiding information overload when activity is exceptionally high.

The Ushahidi ("Ushahidi" n.d.) crowd-reporting platform enables curation and geo-visualization of manually submitted reports from social media sources, email and SMS. To our knowledge, it is the only system specifically designed to handle citizen reports that has been actively used in a large number of real disasters. Due to reliance on users in all information-processing stages, Ushahidi's effectiveness depends entirely on the size, coordination and motivation of crowds. The majority of the most successful deployments have been by the Standby Task Force ("Introducing

the Standby Task Force" n.d.), a volunteer organisation aiming to bring together skilled individuals to remotely provide help in disaster cases, using Internet technologies. For instance, Standby Task Force has set up dedicated teams for media monitoring, translation, verification, and geolocation. This approach adapts well to needs of specific disasters, but it has proven difficult to scale processing capacity to match information inflow rates during the largest events, as was shown during the Haiti earthquake disaster ("Ushahidi Haiti Project Evaluation Final Report" n.d.).

Cameron et al. (2012) developed a system that captures location and volume of Twitter data, providing near real-time keyword search. Their system relies on a trained classifier to detect specific event types, and uses a burst detection method to provide emergency management staff with clues. Twitcident (Abel et al. 2012) is a related Twitter filtering and analysis system that improves situation awareness during small-scale crisis response, such as music festivals and factory fires. It employs classification algorithms to extract messages about very specific events, but is not built to monitor large and complex events with multiple parallel storylines. Both these systems work only with geotagged tweets, which make up around 1 % of all posted messages as of 2013.

Twitcident and the work by Cameron et al. exemplify how despite extensive research into automated classifiers for short contextual strings, classification and information extraction has proven to be significantly harder than for well-formed news articles and blog posts. Like in both of these systems, classifiers tend to be language specific and new training data is needed for each new desired label. This greatly restricts their use in the mass disaster space, where report language is not known beforehand and new report types may be sought in each new disaster.

EMM NewsBrief (n.d.) and Steinberger et al. (2013) automatically mines and clusters mainstream news media from predetermined sources in a wide range of languages, with new summaries produced every 10 min. It too relies on rule-based classifiers for meta-data, but substantial investment has been made to create such rules over a decade. Despite this great investment, it has not been extended to handle social media.

Inspired by the above system, CrisisTracker (Rogstadius et al. 2013b) was developed to enable timely use of social media as a structured information source during mass disasters. Its approach to accomplish this is by combining language-independent fast and scalable algorithms for data collection and event detection, with accurate and adaptable crowd curation. Rather than displaying only high-level statistical metrics (e.g., word clouds and line graphs) and provide search for single social media messages, CrisisTracker's clustering provides event detection, content ranking and summarization while retaining drill-down functionality to raw reports. The system is intended for use during mass disaster and conflict when organizations lack resources to fully monitor events on the ground, or when physical access to local communities is for some reason restricted.
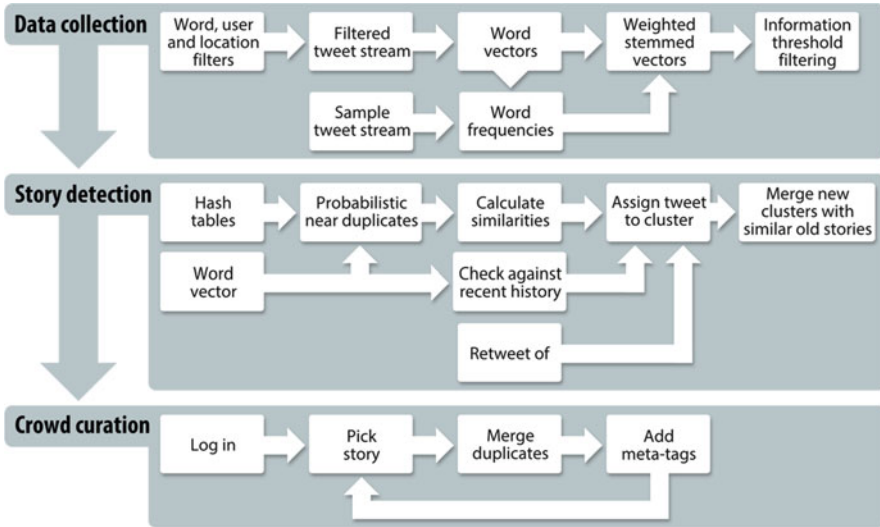
Fig. 4.1 Information processing pipeline in CrisisTracker (Rogstadius et al. 2013b)

4.3.4   Case Study: CrisisTracker

This section provides a summary of how real-time social media mining is conducted by CrisisTracker's information processing pipeline (Fig. 4.1). It consists of data collection, story detection, crowd curation and information consumption. Crowd curation is made possible by decoupling the information itself (stories) from how it has been shared in the social network (tweets). Tweets are collected through Twitter's stream API. This allows a system administrator to define filters in the form of words, geographic bounding boxes and user accounts for which all new matching tweets will be returned as a stream. Generally around 1 % of all tweets are geotagged, thus good keyword filters are the primary way to efficiently obtain information about a topic. Many tweets contain very little information and therefore the system discards messages having fewer than two words after stop word removal and a very low sum of global word weights (approximated inverse document frequencies).

4.3.4.1   Story Detection

Incoming tweets are compared to previously collected tweets using a bag-of-words approach and cosine similarity metric, to group together (cluster) messages that are highly similar. The system uses an extended version of a clustering algorithm for Twitter (Petrovic et al. 2010) based on Locality Sensitive Hashing (Charikar 2002), a probabilistic hashing technique that quickly detects near-duplicates in a stream of

feature vectors. Petrovic et al. (2010) used an initial computation pass to calculate global word statistics (inverse document frequencies) in their offline corpus. In an online setting, word frequencies cannot be assumed to be constant over time, e.g. due to local changes in the tracked event and global activity in different time zones. The algorithm was therefore extended for use in CrisisTracker. Most notably, word statistics are collected based on both the filtered stream and Twitter's sample stream, i.e. a 1 % sample of all posted tweets. For a more detailed explanation on how the tweets are clustered to reflect crisis events in realtime we refer the reader to (Rogstadius et al. 2013b).

CrisisTracker's underlying algorithm offers high precision, but the set of tweets that discuss a particular topic is often split across several clusters. All new clusters are therefore compared with the current clusters to check for overlap. This cluster of clusters is called a story, and this method also enables human intervention in the clustering process. Finally, as the system would quickly run out of storage space if all content was kept, increasingly larger stories and all their content are deleted with increasing age, unless they have been tagged by a human. Stories consisting of a single tweet are kept for approximately 1 day.

4.3.4.2   Crowd Curation and Meta-Data Creation

The reason CrisisTracker clusters the tweet stream into stories is to facilitate crowd curation. De-duplication (ideally) eliminates redundant work, directly reduces the number of items to process per time unit, enables size-based ranking of stories, and groups together reports that mention the same event but contain different details necessary for piecing together a complete narrative.

Search and filtering requires meta-data for stories. Some of this meta-data is extracted automatically, i.e. time of the event (timestamp of first tweet), keywords, popular versions of the report, and number of unique users who mention the story (it's "size"). Story size enables CrisisTracker to estimate how important the message is to the community that has shared it (Rogstadius et al. 2011b). Users of the system can rank stories by their size among all Twitter users, or among the 5000 users most frequently tweeting about the disaster. Typically the top 5000 option better highlights stories with detailed incremental updates to the situation, while the full rank more frequently includes summary articles, jokes and opinions. Since meta-data is assigned per-story, it also covers future tweets in the same story.

Curators are directed towards recent and extensively shared stories, but can self-select which stories to work on. The first curation step is to further improve the clustering, by optionally merging the story with possible duplicate stories that are textually similar but fall below the threshold for automated merging. Miss-classified content can also be removed from stories, which are then annotated (Fig. 4.2) with location, deployment-specific report categories (e.g., infrastructure damage or violence) and named entities. Stories deemed irrelevant (e.g., a recipe named after a location) can be hidden, which prevents them from showing up in search results. Only a Twitter account is required to volunteer as a curator.
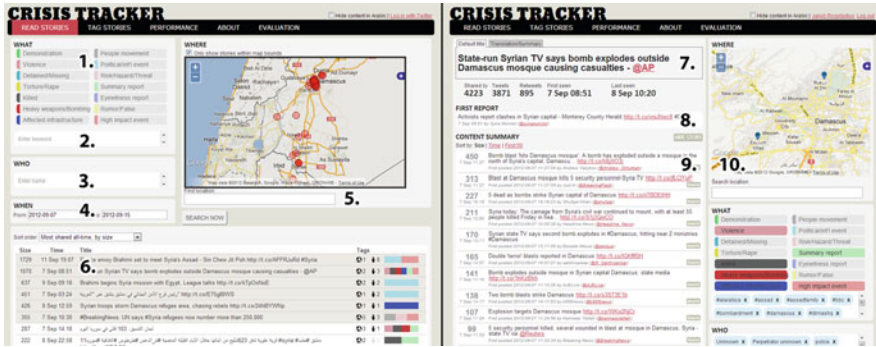
Fig. 4.2 *Left*: User interface for exploring stories, with filters for category (1), keywords (2), named entities (3), time (4) and location (5), with matching stories below (6). *Right*: A single story, with title (7), first tweet (8), grouped alternate versions (9) and human-curated tags (10)

4.3.4.3 Information Consumption

Disaster responders and others interested in the information can filter stories by time, location, report category and named entities. Disaster managers have pointed out (Rogstadius et al. 2013a) that these are basic dimensions along which information is structured in the disaster space. They match how responsibilities are typically assigned within the responder command structure, i.e. by location and/or type of event or intervention. Figure 4.2 presents the interfaces for exploring stories and for reading and curating a single story. The interface for curators to select work items is not shown.

4.3.4.4 CrisisTracker in Action

While during testing and development CrisisTracker was used to monitor events such as Fukushima nuclear disaster and various crises in Middle East, its most large-scale field trial dealt with the 2012 civil war in Syria. In the trial 48 expert curators with prior experience on working with humanitarian disasters signed up to use CrisisTracker as part of their information management toolkit. During the 8-day study CrisisTracker processed 446 000 tweets daily, on average, and managed to successfully reduce the information into consumable stories, thus helping the volunteer curators' tasks. As for concrete findings, CrisisTracker was found successful in enhancing situational awareness of such disaster areas. In practice, it took about 30 min after an isolated incident to happen before CrisisTracker could reduce the social media information overload into a consumable story. This is somewhere between direct eyewitness reports and mass media coverage. CrisisTracker is not, however, a tool to replace existing information management tools. As a research project, it still had its intended impact, as for example certain organisations of UN have specifically requested system features that CrisisTracker

pioneered in this domain. Further details of the mentioned field trial are reported in (Rogstadius et al. 2013b).

4.4   Offline Human Sensors

So far our chapter has focused on collecting data from users, workers, or volunteers who typically sit in front of their desktop computer or who carry a mobile device. In this section we focus on *offline human sensors,* i.e. urban and in-situ systems that collect data from people beyond the desktop environment. We provide an overview of crowdsourcing beyond the desktop, and of systems that are designed to collect opinions from pedestrians in an urban context. We present a case study on a set of systems that use public displays to collect feedback from citizens, and provide strategies and guidelines for conducting this kind of work.

4.4.1   Crowdsourcing Beyond the Desktop

Crowdsourcing with ubiquitous technologies beyond the desktop is increasingly gaining researchers' attention (Vukovic et al. 2010), especially using mobile phones. Similarly as with online crowdsourcing, collecting on-demand information from users on the go practically allows transforming the users into human sensors, capable of providing rich type of feedback about their immediate surroundings as well as about many types of arbitrary issues.

Several mobile platforms for crowdsourcing have been suggested in academia, and quite a few exist as public and fully functional applications as well. Targeting low-end mobile phones, txtEagle (Eagle 2009) is a platform for crowdsourcing tasks specific to habitants of developing countries. Similar platforms are MobileWorks (Narula et al. 2011) and mClerk (Gupta et al. 2012) that specifically focus on asking users to convert handwritten words to typed text from a variety of vestigial dialects. Targeting smartphones, Alt et al. (2010) explore location-based crowdsourcing for distributing tasks to workers. They focus on how workers may actively perform real-world tasks for others, such as giving a real-time recommendation for a restaurant, or providing an instant weather report wherever they are. Similarly, Väätäjä et al. (2011) report a location-aware crowdsourcing platform for authoring news articles by requesting photographs or videos of certain events from its workers. Mashhadi and Capra (2011) suggest using contextual information, such as mobility, as a mechanism to ensure the quality of crowdsourced work.

A very active community has developed around the topic of crowdsourcing measurements and sensing. This participatory sensing movement is also referred to as "Citizen Science" (Paulos et al. 2008) and relies on mobilizing large parts of the population to contribute to scientific challenges via crowdsourcing. Often this involves the use of mobile phones for collecting data (Burke et al. 2006; Goncalves

et al. n.d.) or even donating computational resources while the phone is idle (Arslan et al. 2012).

Despite the appeal of mobile phones, using them for crowdsourcing requires workers' *implicit* deployment, configuration and use of the device. For example, in SMS-based crowdsourcing, participants need to explicitly sign up for the service, at the cost of a text message exchange. This challenges recruitment of workers, as a number of steps need to be performed before a worker can actually start contributing using their device. For these reasons, public displays crowdsourcing has gained popularity recently, since it does not require any deployment effort from the worker to contribute.

A number of previous studies have investigated the use of public interactive displays for the purpose of collecting data, most often collecting explicit human input (Ananny and Strohecker 2009; Brignull and Rogers 2003; Hosio et al. 2012).

Opinionizer (Brignull and Rogers 2003) is a system designed and placed in two authentic social gatherings (parties) to encourage socialization and interaction. Participants could add comments to a publicly visible and shared display. During the study the authors found that a major deterrent preventing people from participating is *social embarrassment*, and suggest making the public interaction purposeful. The environment, both on and around the display, also affect the use and data collected, as the environment produces strong physical and social affordances which people can easily and unambiguously pick up on. Hence they argue for facilitating the public in its needs to rapidly develop their conceptions of the purpose of the social activity, and to be able to move seamlessly and comfortably between being an onlooker and a participant.

A further study that considered public displays as data collection mechanisms was *TextTales* (Ananny and Strohecker 2009). Here the authors attempted to explore the connection between story authorship and civic discourse by installing a large, city-scale, interactive public installation that displays a 3-by-3 grid of image-text combinations. A discussion on a certain photograph would start with SMSs sent by users, displayed in a comments stream. The comments of TexTales users deviated significantly from the "intended" topic of discourse, i.e., the theme set by the photographs. More importantly, this study highlights the challenges in harnessing the general public in natural usage settings for a tightly knit purpose.

Literature suggests that people are interested to use public display deployments (Ananny and Strohecker 2009; Brignull and Rogers 2003; Hosio et al. 2012), but with personal motives in mind resulting in strong appropriation of the technology. For these reasons, a recent study (Goncalves et al. 2013a) was the first attempt to investigate altruistic use of interactive public displays in natural usage settings as a crowdsourcing mechanism. They contrasted a non-paid crowdsourcing service on public displays against the same task being done on a Mechanical Turk (Rogstadius et al. 2011a). The results show that altruistic use, such as for crowdsourcing, is feasible on public displays, and through the controlled use of motivational design and validation check mechanisms, workers' performance can be improved.

An important difference between online crowdsourcing markets and public displays crowdsourcing is the need to login. The login mechanism on Amazon's

Mechanical Turk is a form of quality control that denies access to tasks for workers who perform poorly or attempt to cheat (Mashhadi and Capra 2011). This additional barrier is not necessary on a public display as "bad" workers have no monetary incentive to lose time trying to cheat the system. In this case, potential workers could just approach the public display and start performing tasks right away, instead of going through an authentication mechanism that would most likely greatly diminish the amount of answers gathered.

Finally, Amazon's Mechanical Turk finds it challenging to recruit workers that speak a particular language or live in a particular city (Paolacci et al. 2010). The strategic placement of public displays could help mitigate this issue by, for example, going directly to people that speak a specific language. Another example in which public displays could be used to improve crowdsourcing capabilities would be to target a specific audience with specialized skills that might be difficult to reach otherwise. For example by placing a medical crowdsourcing task (such as the one presented in this paper) on public displays located on a medical school campus it would be possible to reach users at the exact moment when they have free time to do the tasks. In general, it seems that public displays are a highly promising medium to tap into citizens' free time and collecting the public opinion (Hosio et al. 2014).

4.4.2 Collecting Citizen Opinions

Public display research has focused heavily on interaction, attention, and design, but relatively little attention is given to civic engagement. Civic engagement calls for understanding of functional feedback mechanisms. Previously, public displays have been proposed especially as a viable *opportunistic feedback medium* because they allow passersby to understand situated and contextually relevant information, leading to genuinely insightful feedback (Battino Viterbo et al. 2011). Supporting this, Ananny argued that public opinions are highly situated (Ananny and Strohecker, 2009) and De Cindio observed that people leave feedback often during so called *peak* or *protest moments,* when the circumstances for public discourse or disapproval are right (De Cindio et al. 2008). These results together raise the question whether situated feedback mediums could be leveraged to reach people during these key moments for discourse.

One may expect these moments to occur when citizens confront a public display in a city and are given the possibility to leave instant feedback about a locally remarkable and topical issue that invades their territory. Public displays also foster sociality and group use by nature (Kuikkaniemi et al. 2011; Peltonen et al. 2008), and getting feedback from groups of users is often easier than from individuals (Hosio et al. 2012). Furthermore, the well-known *honeypot effect*, referring to the phenomenon of people becoming interested in a display after a single individual first is seen interacting with it (Brignull and Rogers 2003), can be leveraged to our advantage in spreading awareness about the feedback channel among nearby potential users.

Archetypal feedback applications on public displays utilize typing in some form as their main input modality. Earlier, for example, Twitter has been trialed as an input mechanism for public displays. The experiments with *Discussions In Space* (Schroeter et al. 2012) highlighted especially how content about the display location itself work well for engaging audiences, and how the interfaces in uncontrolled environments must be self-explanatory and offer clear cues to users on how they can participate. Ananny and Strohecker leveraged public screens and SMS to create public opinion forums (Ananny and Strohecker 2009). Their *TexTales* installations highlighted how urban spaces can become sites for collective expression and nurture informal, often amusing discussions among its habitants.

A playful feedback application, connected to social networking services and utilizing a virtual keyboard and a web camera for feedback was introduced by Hosio et al. (2012). Studies with *Ubinion* also highlighted situated public displays being fit for acquiring contextually relevant feedback. Similar projects (Day et al. 2007; Munson et al. 2011) developed feedback systems for campus settings, utilizing online interfaces, dedicated mobile clients, and Twitter as inputs. In these studies, Twitter was suggested as a good tool to provide content for public displays, and SMS was envisioned handier for feedback than dedicated mobile applications.

4.4.3   Case Study: Opinions for Civic Engagement

We present a case study where public displays were used as a mechanism for collecting civic feedback. This was prompted by a major renovation of the city centre, which included building new pavement and underground heating systems for two of the busiest pedestrian streets in downtown Oulu, Finland. This heavily affected pedestrian flows and everyday business in all the surrounding areas, and was a heated topic in this city, and it was reported in dozens of stories in local newspapers, where it garnered heavy attention in the discussion sections both for and against the project.

In this case study, the displays used were 57″ full-HD touch screen displays with rich connectivity options, fitted in weather-proof casings. Many of the displays had been located in the vicinity of the renovation area already for several years and as such have gone beyond novelty to be an accepted part of the city infrastructure itself (Ojala et al. 2012a). The displays were on either end of each of the walking streets and one at their crossing, making them situated close to the project. Besides these five displays, at all times there were three to six more displays located elsewhere in downtown and other pivotal public spaces in the city. Figure 4.3 depicts the renovation environment and one of the displays next to the renovation area.

The tested system was an application for the public displays that allowed citizens to rate the progress of the renovation, and to provide open-ended feedback. The application was available to any member of the public in a 24/7 fashion on all displays. Civic engagement should be made available to all social groups (Mohammadi et al. 2011). Therefore, studying a system "in the wild" where

Fig. 4.3 From *left*: a conceptual image of how the new renovated street will look like (used with permission from Oulu Technical Centre), a display in the end of the same street, and the actual renovation taking place in Downtown Oulu

everyone can use is a fundamental requirement for these types of systems. This is not always easy, as the urban space itself is a rich but challenging environment to deploy pervasive infrastructure and applications (Müller et al. 2010). Several considerations, including the intertwined social practices of the area, robustness of the technology, abuse, vandalism, balance between the different stakeholders, and even weather conditions may cause constraints when deploying in the wild (Alt et al. 2011; Dalsgaard and Halskov 2010; Greenfield and Shepard 2007; Huang et al. 2007; McCullough 2004). However, to gain an understanding of how technology is received and appropriated by the general public, deployment in authentic environments, or *living laboratories,* is highly beneficial (Rogers et al. 2007; Sharp and Rehman 2005). This type of case study follows Brown's advice (Brown et al. 2011) to move beyond reporting artificial success: rather than proposing a solution that fulfils all the needs of all involved stakeholders, the study can report what happened with the chosen solutions in the complicated setting.

During the 3-month pilot, the application for providing feedback was launched 2664 times by citizens, which resulted in 81 text based feedbacks and 66 sets of likert-scale ratings. Thus, 3.0 % of all application launches led to users leaving textual feedback, and 8.0 % led to users using the smiley based mechanism. This strongly reflects *lurking* behaviour online, where up to 99 % of users do not participate in discussions, but rather follow and read information (Preece et al. 2004). The term *lurker* has an unreasonably bad connotation to it. After all, lurking is in many cases beneficial for the greater community, and a case can be even made for lurking to be normal behaviour and participation abnormal: who would be reading if everybody focused on contributing (Nonnecke and Preece 2000)?

Müller argues that public displays do not invite people for a single reason, but users come across them with no dedicated purpose (Müller et al. 2010). Further, when a display features multiple applications, many application launches are caused by curiosity or play rather than intention of using them (Hosio et al. 2010). These findings together suggest that part of the application launches was not intentional, and that if the applications were deployed on bespoke displays, the participation rate would be higher.

Several factors suggest civic engagement to be challenging. Downs has observed that citizens appear to be "rationally ignorant" of topical issues and local policies, because in their opinion the feedback they give will not be influential (Downs 1957). In this case study, the plans for the renovation were already finished and published, and it was not realistic to affect the final outcome anymore. Another consideration is target demographics. It is only fair to assume that a municipal renovation project concerns a more mature audience, i.e. taxpayers who in the end pay for it. Clary and Snyder (2002) report that it is generally harder to get feedback from an adult audience than from the young, as adults often have deeply grained habits that simply do not support community-driven participation.

However, the results of the case study remain carefully optimistic about the overall participation. While the total of 81 feedback messages (27 relevant) collected may not be a lot—especially when compared to the results of related feedback prototypes in literature—the city authorities reported it was the *only* feedback they ever received from citizens in the course of this case study. Their conventional feedback mechanisms, phone and email, were not used for citizen feedback, and they were overall very satisfied with the performance of the new feedback channel.

4.4.4 Strategies and Guidelines for Eliciting Citizen Feedback

Based on the case study described above, as well as literature, certain recommendations are presented for researchers planning to orchestrate longitudinal studies in civic engagement with public displays. First, one should expect *social use* of this technology. Social and performative uses are intrinsic factors that drive the use of public displays (Kuikkaniemi et al. 2011; O'Hara et al. 2008; Ojala et al. 2012b; Peltonen et al. 2008). This has to be considered when designing feedback applications by cultivating social use, not by trying to steer away from it. For example, Brignull and Rogers (2003) findings suggest an awkwardness and social pressure that people feel when interacting alone with public displays. Third-party stakeholders should be educated about this already early in the design phase of civic engagement installations. Hence, it is suggested to *avoid topics of civic discourse that call for participation by individuals*.

One should also s*et realistic goals* for this kind of research. It is established that various social needs, such as self-expression or ill-behaviour, present themselves in the use of new communication channels (Harper 2010; Kindberg et al. 2005; Van House 2007). If a feedback channel is deployed in the wild and allows free form submission, these needs are likely to lead to appropriation, i.e. increased amount of off-topic feedback.

Studying several related feedback applications often leads to believing that getting tens of even hundreds of feedback messages with just a few installations is technically easy (Ananny and Strohecker 2009; Battino Viterbo et al. 2011; Brignull and Rogers 2003; Hosio et al. 2012). However, common in all these prototypes is informal or amusing topics of feedback and discussion. Civic engagement, on the

contrary, often has to do with narrow, predefined topic of interest to a given local authority. As such, it lacks mass-appeal (Uslaner and Brown 2005). Further, people are ignorant towards civic engagement (Downs 1957), and habits of especially adults do not support participation (Clary and Snyder 2002). When conducting research in unsupervised environments and with uncoached users, it is important to acknowledge that the participation rate may deteriorate rapidly.

It is true that perhaps controlled, situated trials could be used to elicit the same amount of feedback that this installation was capable of doing. However, sustained participation calls for longitudinal action, according to Clary and Snyder (2002), and has other benefits too. Due to its opportunistic nature, it will reach users that would otherwise be truly unreachable, as demonstrated successfully in (Hosio et al. 2012), where 67 % of the public display users had not been connected before with the corresponding authorities. The social settings, target audience, used feedback mechanisms, and the feedback topic of civic engagement all play a role in the actual and argued success of a deployment. With too high initial expectations, it will be hard to judge success later on. Hence, an important recommendation is to *be aware of what is realistic participation in a deployment in uncontrolled and authentic settings.*

4.5 Conclusion

This chapter provides an overview and multiple case studies of systems where humans are the primary source of information. Recent technological advances in making communication more affordable, computation faster, but also the changing norms regarding use of technology, have enabled a range of new applications and systems that collect data from humans. We have described how online crowdsourcing markets are enabling the collection of data from humans in a systematic way, and how harvesting of online social media can offer real-time insights into evolving events. We also provide an overview of interactive urban technologies that collect data from pedestrians in-situ.

References

Abel, F., Hauff, C., Houben, G.-J., Stronkman, R., Tao, K.: Semantics + filtering + search = twitcident. Exploring information in social web streams. In: HT'12: Proceedings of the 23rd ACM Conference on Hypertext and Social Media, pp. 285–294. ACM (2012). doi:10.1145/2309996.2310043

Alt, F., Memarovic, N., Elhart, I., Bial, D., Schmidt, A., Langheinrich, M., Harboe, G., Huang, E., Scipioni, M.P.: Designing shared public display networks: implications from today's paper-based notice areas. In: Pervasive'11: Proceedings of the 9th International Conference on Pervasive Computing, pp. 258–275. Springer-Verlag, Berlin, Heidelberg (2011). doi:10.1007/978-3-642-21726-5_17

Alt, F., Shirazi, A.S., Schmidt, A., Kramer, U., Nawaz, Z.: Location-based crowdsourcing: extending crowdsourcing to the real world. In: NordiCHI'10: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 13–22. ACM (2010). doi:10.1145/1868914.1868921

Ananny, M., Strohecker, C.: TextTales: creating interactive forums with urban publics. In: Foth, M. (ed.) Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City. IGI Global, Hershey, PA (2009). doi:10.4018/978-1-60566-152-0.ch005

Arslan, M.Y., Singh, I., Singh, S., Madhyastha, H.V., Sundaresan, K., Krishnamurthy, S.V.: Computing while charging: Building a distributed computing infrastructure using smartphones. In: CoNEXT'12: Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies, pp. 193–204. ACM (2012). doi:10.1145/2413176.2413199

Battino Viterbo, P., Barsotti, V., Vande Moere, A.: Situated & social feedback in the city. In: Workshop on Large Displays in Urban Life from Exhibition Halls to Media Facades (2011)

Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., Kraut, R.E.: Using social psychology to motivate contributions to online communities. In: CSCW'04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 212–221. ACM (2004). doi:10.1145/1031607.1031642

Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. J. Comput. Sci. **2**(1), 1–8 (2011). doi:10.1016/j.jocs.2010.12.007

Brignull, H., Rogers, Y.: Enticing people to interact with large public displays in public spaces. In: Proceedings of the INTERACT, vol. 3, pp. 17–24. (2003)

Brown, B., Reeves, S., Sherwood, S.: Into the wild: challenges and opportunities for field trial methods. In: CHI'11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1657–1666. ACM (2011). doi:10.1145/1978942.1979185

Burke, J.A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., Srivastava, M.B.: Participatory sensing. In: First Workshop on World-Sensor-Web: Mobile Device Centric Sensory Networks and Applications at Sensys'06 (2006)

Cameron, M.A., Power, R., Robinson, B., Yin, J.: Emergency situation awareness from twitter for crisis management. In: WWW'12 Companion: Proceedings of the 21st International Conference Companion on World Wide Web, pp. 695–698. ACM (2012). doi:10.1145/2187980.2188183

Chandler, D., Kapelner, A.: Breaking monotony with meaning: motivation in crowdsourcing markets. J. Econ. Behav. Organ. **90**, 123–133 (2013). doi:10.1016/j.jebo.2013.03.003

Charikar, M.S.: Similarity estimation techniques from rounding algorithms. In: STOC'02: Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing, pp. 380–388. ACM (2002). doi:10.1145/509907.509965

De Cindio, F., Di Loreto, I., Peraboni, C.A.: Moments and modes for triggering civic participation at the urban level. In: Foth, M. (ed.) Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City, pp. 97–113. IGI Global, Hershey, PA (2008). doi:10.4018/978-1-60566-152-0.ch007

Clary, E.G., Snyder, M.: Community involvement: opportunities and challenges in socializing adults to participate in society. J. Soc. Issues **58**(3), 581–591 (2002). doi:10.1111/1540-4560.00277

Cosley, D., Frankowski, D., Kiesler, S., Terveen, L., Riedl, J.: How oversight improves member-maintained communities. In: CHI'05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 11–20. ACM (2005). doi:10.1145/1054972.1054975

Dalsgaard, P., Halskov, K.: Designing urban media facades: cases and challenges. In CHI'10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2277–2286. ACM (2010). doi:10.1145/1753326.1753670

Day, N., Sas, C., Dix, A., Toma, M., Bevan, C., Clare, D.: Breaking the campus bubble: informed, engaged, connected. In: BCS-HCI'07: Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI . . . But Not as We Know It, vol. 2, pp. 133–136. British Computer Society (2007)

Deci, E.: Intrinsic Motivation. Plenum, New York (1975)

Downs, A.: An economic theory of democracy. In: An Economic Theory of Democracy. Harper & Row, New York (1957)

Dugas, A.F., Hsieh, Y.-H., Levin, S.R., Pines, J.M., Mareiniss, D.P., Mohareb, A., Gaydos, C.A., Perl, T.M., Rothman, R.E.: Google flu trends: Correlation with emergency department influenza rates and crowding metrics. Clin. Infect. Dis. **54**(4), 463–469 (2012). doi:10.1093/cid/cir883

Eagle, N.: Txteagle: mobile crowdsourcing. In: IDGD'09: Proceedings of the 3rd International Conference on Internationalization, Design and Global Development: Held as Part of HCI International 2009, pp. 447–456. Springer, Berlin, Heidelberg (2009). doi:10.1007/978-3-642-02767-3_50

EMM NewsBrief.: EMM newsbrief. [Web page]. Retrieved from http://press.jrc.it (n.d.)

Etzioni, A.: Modern Organizations. Prentice-Hall, Englewood Cliffs, NJ (1971)

Gibbons, R.: Incentives and careers in organizations. In: Kreps, D., Wallis, K. (eds.) Advances in economic theory and econometrics, vol. II, pp. 1–37. Cambridge University Press, Cambridge (1997)

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L.: Detecting influenza epidemics using search engine query data. Nature **457**(7232), 1012–1014 (2009). doi:10.1038/nature07634

Gneezy, U., Rustichini, A.: Pay enough or don't pay at all. Q. J. Econ. **115**(3), 791–810 (2000). doi:10.1162/003355300554917

Goel, S., Hofman, J.M., Lahaie, S., Pennock, D.M., Watts, D.J.: Predicting consumer behavior with web search. Proc. Natl. Acad. Sci. **107**(41), 17486–17490 (2010). doi:10.1073/pnas.1005962107

Goncalves, J., Ferreira, D., Hosio, S., Liu, Y., Rogstadius, J., Kukka, H., Kostakos, V.: Crowdsourcing on the spot: Altruistic use of public displays, feasibility, performance, and behaviours. In: UbiComp'13: Ubicomp, pp. 753–762. ACM (2013). doi:10.1145/2493432.2493481

Goncalves, J., Kostakos, V., Hosio, S., Karapanos, E., Lyra, O.: IncluCity: Using contextual cues to raise awareness on environmental accessibility. In: ASSETS'13: ASSETS, pp. 17:1–17:8. ACM (2013). doi:10.1145/2513383.2517030

Goncalves, J., Kostakos, V., Karapanos, E., Barreto, M., Camacho, T., Tomasic, A., Zimmerman, J.: Citizen motivation on the go: the role of psychological empowerment. Inter. Comput. (n.d.) doi:10.1093/iwc/iwt035

Greenfield, A., Shepard, M.: Situated Technologies Pamphlets 1: Urban Computing and Its Discontents. The Architectural League of New York, New York (2007)

Gupta, A., Thies, W., Cutrell, E., Balakrishnan, R.: MClerk: enabling mobile crowdsourcing in developing regions. In: CHI'12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1843–1852. ACM (2012). doi:10.1145/2207676.2208320

Harper, F.M., Raban, D., Rafaeli, S., Konstan, J.A.: Predictors of answer quality in online Q&A sites. In: CHI'08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 865–874. ACM (2008). doi:10.1145/1357054.1357191

Harper, R.: Texture: Human Expression in the Age of Communications Overload. MIT, Cambridge (2010)

Heer, J., Bostock, M.: Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In: CHI'10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 203–212. ACM (2010). doi:10.1145/1753326.1753357

Horton, J.J., Rand, D.G., Zeckhauser, R.J.: The online laboratory: conducting experiments in a real labor market. Exp. Econ. **14**(3), 399–425 (2011). doi:10.1007/s10683-011-9273-9

Hosio, S., Goncalves, J., Kostakos, V., Riekki, J.: Exploring civic engagement on public displays. In: User-Centric Technology Design for Nonprofit and Civic Engagements, pp. 91–1119. Springer (2014). doi:10.1007/978-3-319-05963-1_7

Hosio, S., Kostakos, V., Kukka, H., Jurmu, M., Riekki, J., Ojala, T.: From school food to skate parks in a few clicks: using public displays to bootstrap civic engagement of the young. In: Pervasive, pp. 425–442. Springer (2012). doi:10.1007/978-3-642-31205-2_26

Hosio, S., Kukka, H., Jurmu, M., Ojala, T., Riekki, J.: Enhancing interactive public displays with social networking services. In: MUM'10: Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia, pp. 23:1–23:9. ACM (2010). doi:10.1145/1899475.1899498

Huang, E.M., Mynatt, E.D., Trimble, J.P.: When design just isn't enough: the unanticipated challenges of the real world for large collaborative displays. Pers. Ubiquit. Comput. **11**(7), 537–547 (2007). doi:10.1007/s00779-006-0114-3

Hubbard, R.G., Palia, D.: Executive pay and performance evidence from the US banking industry. J. Financ. Econ. **39**(1), 105–130 (1995). doi:10.1016/0304-405X(94)00816-J

Introducing the Standby Task Force: Introducing the standby task force. [Web page]. Retrieved from http://blog.standbytaskforce.com/about/introducing-the-standby-task-force/ (n.d.)

Kindberg, T., Spasojevic, M., Fleck, R., Sellen, A.: The ubiquitous camera: an in-depth study of camera phone use. IEEE Pervasive Comput. **4**(2), 42–50 (2005). doi:10.1109/MPRV.2005.42

Kostakos, V., Juntunen, T., Goncalves, J., Hosio, S., Ojala, T.: Where am I? Location archetype keyword extraction from urban mobility patterns. PloS One **8**(5), e63980 (2013). doi:10.1371/journal.pone.0063980

Kuikkaniemi, K., Jacucci, G., Turpeinen, M., Hoggan, E., Müller, J.: From space to stage: how interactive screens will change urban life. Computer **44**(6), 40–47 (2011). doi:10.1109/MC.2011.135

Laput, G., Lasecki, W.S., Wiese, J., Xiao, R., Bigham, J.P., Harrison, C.: Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds. In: CHI'15: Proceedings of the International ACM Conference on Human Factors in Computing Systems, pp. 1935–1944. (2015). doi:10.1145/2702123.2702416

Lazear, E.P.: Performance pay and productivity. Am. Econ. Rev. **90**(5), 1346–1361 (2000). doi:10.1257/aer.90.5.1346

Mashhadi, A.J., Capra, L.: Quality control for real-time ubiquitous crowdsourcing. In: UbiCrowd'11: Proceedings of the 2nd International Workshop on Ubiquitous Crowdsouring, pp. 5–8. ACM (2011). doi:10.1145/2030100.2030103

Mason, W., Watts, D.J.: Financial incentives and the "performance of crowds". In: HCOMP'09: Proceedings of the ACM SIGKDD Workshop on Human Computation, pp. 77–85. ACM (2009). doi:10.1145/1600150.1600175

McCullough, M.: Digital Ground: Architecture, Pervasive Computing, and Environmental Knowing. MIT, Cambridge (2004)

Mohammadi, S.H., Norazizan, S., Shahvandi, A.R.: Civic engagement, citizen participation and quality of governance in iran. J. Hum. Ecol. **36**(3), 211–216 (2011)

Munson, S.A., Rosengren, E., Resnick, P.: Thanks and tweets: comparing two public displays. In: CSCW'11: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, pp. 331–340. ACM (2011). doi:10.1145/1958824.1958875

Müller, J., Alt, F., Michelis, D., Schmidt, A.: (2010). Requirements and design space for interactive public displays. In: MM'10: Proceedings of the International Conference on Multimedia, pp. 1285–1294. ACM. doi:10.1145/1873951.1874203

Narula, P., Gutheim, P., Rolnitzky, D., Kulkarni, A., Hartmann, B.: MobileWorks: a mobile crowdsourcing platform for workers at the bottom of the pyramid. In: AAAI Workshop on Human Computation, pp. 121–123. (2011)

Nguyen, D.H., Marcu, G., Hayes, G.R., Truong, K.N., Scott, J., Langheinrich, M., Roduner, C.: Encountering sensecam: personal recording technologies in everyday life. In: Ubicomp'09: Proceedings of the 11th International Conference on Ubiquitous Computing, pp. 165–174. ACM (2009). doi:10.1145/1620545.1620571

Nonnecke, B., Preece, J.: Lurker demographics: Counting the silent. In: CHI'00: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80. ACM (2000). doi:10.1145/332040.332409

O'Hara, K., Glancy, M., Robertshaw, S.: Understanding collective play in an urban screen game. In: CSCW'08: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, pp. 67–76. ACM (2008). doi:10.1145/1460563.1460576

Ojala, T., Kostakos, V., Kukka, H., Heikkinen, T., Linden, T., Jurmu, M., Hosio, S., Kruger, F., Zanni, D.: Multipurpose interactive public displays in the wild: three years later. Computer 45(5), 42–49 (2012). doi:10.1109/MC.2012.115

Ojala, T., Kruger, F., Kostakos, V., Valkama, V.: Two field trials on the efficiency of unsolicited bluetooth proximity marketing. In: MUM'12: MUM, pp. 38:1–38:4. ACM (2012). doi:10.1145/2406367.2406414

Paolacci, G., Chandler, J., Ipeirotis, P.: Running experiments on amazon mechanical turk. Judgment Decis. Mak. 5(5), 411–419 (2010)

Paulos, E., Honicky, R.J., Hooker, B.: Citizen science: enabling participatory urbanism. In: Foth, M. (ed.) Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City, pp. 414–436. IGI Global, Hershey, PA (2008). doi:10.4018/978-1-60566-152-0.ch028

Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., Saarikko, P.: It's mine, don't touch!: Interactions at a large multi-touch display in a city centre. In: CHI'08: CHI, pp. 1285–1294. (2008). Retrieved from http://portal.acm.org/citation.cfm?id=1357054.1357255

Perttunen, M., Mazhelis, O., Cong, F., Kauppila, M., Leppänen, T., Kantola, J., Collin, J., Pirttikangas, S., Haverinen, J., Ristaniemi, T., Riekki, J.: Distributed road surface condition monitoring using mobile phones. In: Ubiquitous Intelligence and Computing, pp. 64–78. Springer, Berlin Heidelberg (2011). doi:10.1007/978-3-642-23641-9_8

Petrovic, S., Osborne, M., Lavrenko, V.: Streaming first story detection with application to twitter. In: HLT'10: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 181–189. Association for Computational Linguistics, Stroudsburg (2010)

Polgreen, P.M., Chen, Y., Pennock, D.M., Nelson, F.D., Weinstein, R.A.: Using internet searches for influenza surveillance. Clin. Infect. Dis. 47(11), 1443–1448 (2008). doi:10.1086/593098

Preece, J., Nonnecke, B., Andrews, D.: The top five reasons for lurking: improving community experiences for everyone. Comput. Hum. Behav. 20(2), 201–223 (2004). doi:10.1016/j.chb.2003.10.015

Qin, C., Bao, X., Roy Choudhury, R., Nelakuditi, S.: TagSense: A smartphone-based approach to automatic image tagging. In: MobiSys'11: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, pp. 1–14. ACM (2011). doi:10.1145/1999995.1999997

Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R.E., Hursey, J., Toscos, T.: Why it's worth the hassle: the value of in-situ studies when designing Ubicomp. In: UbiComp'07: Proceedings of the 9th International Conference on Ubiquitous Computing, pp. 336–353. Springer, Berlin, Heidelberg (2007). doi:10.1007/978-3-540-74853-3_20

Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., Vukovic, M.: An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets. In: AAAI Conference on Weblogs and Social Media, pp. 321–328. Barcelona, Spain (2011)

Rogstadius, J., Kostakos, V., Laredo, J., Vukovic, M.: A real-time social media aggregation tool: Reflections from five large-scale events. In: ECSCW Workshop CSCWSmart. Aarhus, Denmark (2011)

Rogstadius, J., Teixeira, C., Karapanos, E., Kostakos, V.: An introduction for system developers to volunteer roles in crisis response and recovery. In: International Conference on Information Systems for Crisis Response and Management, pp. 874–883. Baden-Baden, Germany (2013)

Rogstadius, J., Teixeira, C., Vukovic, M., Kostakos, V., Karapanos, E., Laredo, J.: CrisisTracker: crowdsourced social media curation for disaster awareness. IBM J. Res. Dev. 57(5), 4:1–4:13 (2013). doi:10.1147/JRD.2013.2260692

Rui, H., Liu, Y., Whinston, A.: Whose and what chatter matters? The effect of tweets on movie sales. Decis. Support Syst. 55(4), 863–870 (2013). doi:10.1016/j.dss.2012.12.022

Sahana Foundation: Sahana foundation. [Web page]. Retrieved from http://sahanafoundation.org/ (n.d.)

Schroeter, R., Foth, M., Satchell, C.: People, content, location: Sweet spotting urban screens for situated engagement. In: DIS'12: Proceedings of the Designing Interactive Systems Conference, pp. 146–155. ACM (2012). doi:10.1145/2317956.2317980

Sharp, R., Rehman, K.: The 2005 UbiApp workshop: what makes good application-led research? IEEE Pervasive Comput. **4**(3), 80–82 (2005). doi:10.1109/MPRV.2005.69

Steinberger, R., Pouliquen, B., Van Der Goot, E.: An introduction to the Europe media monitor family of applications. (2013). *ArXiv:1309.5290*

Ushahidi: Ushahidi. [Web page]. Retrieved from http://ushahidi.org (n.d.)

Ushahidi Haiti Project Evaluation Final Report: [Web page]. Retrieved from http://blog.ushahidi.com/2011/04/19/ushahidi-haiti-project-evaluation-final-report/

Uslaner, E.M., Brown, M.: Inequality, trust, and civic engagement. Am. Pol. Res. **33**(6), 868–894 (2005). doi:10.1177/1532673X04271903

Van House, N.A.: Flickr and public image-sharing: Distant closeness and photo exhibition. In: CHI EA'07: CHI'07 Extended Abstracts on Human Factors in Computing Systems, pp. 2717–2722. ACM (2007). doi:10.1145/1240866.1241068

Väätäjä, H., Vainio, T., Sirkkunen, E., Salo, K.: Crowdsourced news reporting: Supporting news content creation with mobile phones. In: MobileHCI'11: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, pp. 435–444. ACM (2011). doi:10.1145/2037373.2037438

VirtualAgility WorkCenter: VirtualAgility workcenter. [Web page]. Retrieved from http://www.virtualagility.com/ (n.d.)

Vukovic, M., Kumara, S., Greenshpan, O. Ubiquitous crowdsourcing. In: Ubicomp'10 Adjunct: Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing—Adjunct, pp. 523–526. ACM (2010). doi:10.1145/1864431.1864504

Wheelmap: Wheelmap. [Web page]. Retrieved from http://wheelmap.org/ (n.d.)