A

Ohmage: A General and Extensible End-to-End Participatory Sensing Platform

H. Tangmunarunkit, and C.K. Hsieh, UCLA
J. Jenkins, C. Ketcham, and J. Selsky, Cornell Tech
F. Alquaddoomi, J. Kang, Z. Khalapyan, B. Longstaff, S. Nolen, J. Ooms, and N. Ramanathan, UCLA
D. Estrin, Cornell Tech

Participatory Sensing is a distributed data collection and analysis approach where individuals, acting alone or in groups, use their personal mobile devices to systematically explore interesting aspects of their lives and communities [Burke et al. 2006]. These mobile devices can be used to capture diverse spatio-temporal data through both intermittent self-report and continuous recording from on-board sensors and applications.

ohmage (http://ohmage.org) is a general, modular and extensible open-source, mobile to web participatory sensing platform that records, stores, analyzes, and visualizes data from both prompted self-report and continuous data streams. These data streams are scriptable and configurable and can be dynamically deployed in diverse settings. Ohmage provides a rich set of end-to-end modules, spanning mobile, server, and web layers. Feedback from hundreds of behavioral and technology researchers, focus group participants, and end-users has been integrated into ohmage through an iterative design process. Ohmage has been used as an enabling platform in over 20 independent projects in many disciplines. We present the flexibility, modularity, and extensibility of ohmage in supporting diverse deployment settings through three distinct case studies in education, health, and clinical research.

Categories and Subject Descriptors: H.4 [**Information Systems**]: Information Systems Applications

General Terms: Design, Platform, Experiences, Deployments, Education, Health, Mobile Devices, Citizen Science, Data Collection

Additional Key Words and Phrases: Participatory Sensing, Deployment Experiences, Instructional Tool, Experience Sampling, Mobile Sensing, Clinical Research

1. INTRODUCTION

Advances in mobile technology, and the ubiquity of mobile phones have enabled phones to become a convenient, affordable and scalable real-time data collection platform. Participatory Sensing (PS) is a distributed data collection and analysis approach that takes advantage of pervasive smartphones [Burke et al. 2006]. In a PS project, participants use their personal phones to systematically collect data relevant to themselves or their community.

PS can facilitate and enhance many different applications: In health, it can assist individuals and their care providers in monitoring and managing symptoms, side effects, and treatments for chronic illness outside the clinical setting [Chen et al. 2012; Ramanathan et al. 2013; Krishna et al. 2009]. In emergency response scenarios, it can be used by citizens or emergency response teams to rapidly gather information in the field [Starbird and Palen 2011; Okolloh 2009]. In education, it can be used as an innovative instructional tool for teaching data collection and analysis concepts and methods [Mobilize 2010; Heggen et al. 2012]. All these scenarios share in-situ data collection, centralized data storage and management, and the need for deployment administration, data analysis, and visualization. However, construction of such a PS system is not the main objective of any of these projects. Development of such a system from scratch would draw time and resources away from their primary activities. There is a clear value to be gained from a general, reusable platform to help construct PS deployments.

Some of the most influential PS platforms preceded development of ohmage. Ushahidi [Okolloh 2009] targeted feature phone users with its primary modality being SMS text messaging. OpenDataKit [Hartung et al. 2010] was designed for semi-professional data collection tasks such as fieldworkers. Ohmage was designed specifically to allow easy configuration of systematic data collection projects in which individuals capture data in the course of their everyday lives using and leveraging the advanced features of smartphones. In this way it is closest to Intel's MyExperience [Froehlich et al. 2007], but with a modular system architecture focused on supporting broader deployment settings and research disciplines.

Ohmage is designed to be a general, modular, and extensible PS platform. Ohmage not only supports scriptable survey-based data, as in OpenDataKit and MyExperience, but also supports scriptable passive data collected from sensors or applications on mobile devices. Ohmage end-to-end functionality is provided through a rich, robust software suite that supports data collection, storage, management, monitoring, analysis, and visualization. At the core of the software suite is the ohmage backend that provides secure web-based APIs for data access. These APIs allow different ohmage components and custom applications to be independently developed and integrated. In addition, ohmage supports fine-grained access controls allowing different users different access privileges depending on their roles in the deployment. Also, ohmage group management supports isolation between different deployments, allowing them to coexist on the same backend. Ohmage key system designs and architecture are provided in Section 2 and 3, respectively.

The ohmage software development methodology is based on an iterative design process that incorporated feedback from researchers, focus groups, and end-users. During the course of its development and evolution, ohmage has been used as an enabling platform in a number of independent behavioral, health, wellness and education projects addressing different populations (e.g., breast cancer survivors, new moms, HIV+ men, young adults with attention deficit hyperactivity disorder (ADHD), etc). Of these, we explore in more detail in Section 4 three unique case studies which demonstrate the breadth and flexibility of ohmage in supporting different research disciplines. These studies are: 1) Mobilize: the ohmage end-to-end PS process is used for teaching computational thinking in 81 Math, Science, and Computer Science high-school classrooms to-date; 2) Moms: 45 young moms used ohmage for 6 months to study diet, stress, and exercise behaviors in order to better manage risk factors related to heart disease; and 3) PREEMPT: ohmage will be used to setup and monitor personalized randomized control trials (N-of-1) of pain treatments in over 100 participants with chronic pain.

2. OHMAGE SYSTEM DESIGN

Prior to ohmage, we developed multiple siloed participatory sensing systems, each specifically implemented for an individual study [CENS 2006; Mun et al. 2009; Acker et al. 2010]. This approach discouraged code reuse, incurred high development and maintenance overhead, and did not scale well. These drawbacks motivated the development of a general and extensible system, ohmage. Since its inception in 2010, ohmage has adopted an iterative design methodology. As ohmage evolves, different versions have been periodically released to support multiple studies in different research areas including behavioral, health, wellness and education. Feedback from our collaborators, researchers, focus groups, and participants; study requirements; as well as lessons learned from past deployments have been continuously and incrementally prioritized and incorporated into ohmage in each release. Key system design criteria that have guided the ohmage development are summarized below.

— **Scriptable data streams** to support diverse data collection requirements. The system should allow users to define the data streams that they want to collect, and automatically deploy the descriptions to participants.

— **Highly available apps.** The mobile apps should allow users to collect data anywhere and any time with or without network connectivity. The apps should also be available on a wide-range of mobile devices to enable more users to use their personal mobile devices to collect data, instead of carrying an extra device. The ohmage self-report apps currently support iOS and Android devices, which together account for 93.9% of smartphone market share in Q3 2013 [comScore 2013].
— **Mobile app usability.** In addition to ease of use, the mobile apps should provide mechanisms to support user engagement. Examples of these mechanisms are reminders to prompt the user to take surveys that fit their schedule, home-screen buttons for quick data entry, and visual feedback to encourage user participation. Most users found the ohmage apps to be easy to use (see Section 4). The reminders, buttons, and feedback are available on our Android app.
— **Study manageability.** The system should provide an interface for researchers to manage their studies, monitor the incoming data, and visualize study analytics. For example, a researcher should be able to examine the data quality and rectify any deployment problems that have arisen.
— **Configurable and customizable.** The system should be configurable to support diverse deployment scenarios. For example, different deployments need different features (e.g. no self-registration), default settings (e.g. max data size), branding (e.g. logo), and terminology. Ohmage provides mechanisms to change system-wide settings, as well as manage accounts and access controls to fit each deployment.
— **Modular and extensible.** The system should allow modules or functionality to be changed or added independently without impacting the rest of the system. Ohmage provides a well-defined set of Application Programming Interfaces (APIs), that are backward compatible within major releases, to securely and centrally create and manage ohmage objects. Through the use of these APIs, different applications (e.g. new mobile apps, customized analysis/visualization modules) can be independently created to connect to the ohmage backend without affecting other applications. Additionally, ohmage provides a rich set of end-to-end PS modules for different scenarios.

3. OHMAGE SYSTEM ARCHITECTURE

The ohmage architecture, shown in Figure 1, consists of four primary components: 1) Ohmage backend that serves as a datastore and provides a unified interface for data access; 2) Mobile data collection apps that run on participants' phones for data collection and feedback; 3) Web-based data management and administration tools for study management and administration; and 4) Web-based data analysis and visualization tools for exploring, analyzing and visualizing the collected data. In this section, we first describe the supported data streams, followed by the detail of each component.

3.1. Ohmage Data Streams

Ohmage supports two types of data streams: 1) self-report data, and 2) passive data. We describe these data streams and how they are configured below.

Self-report data

Self-report data refers to data captured by a prompted experience sampling method where participants are asked to record their observations and experiences in the form of surveys [Froehlich et al. 2007]. A *project* describes the structure and properties of the self-report data to be collected. A project contains a set of surveys, each of which is a sequence of messages or prompts to be displayed. A message (e.g. "Please get up and walk for 2 minutes.") is a brief communication which does not require user input. A prompt solicits a response from the user; either an answer to a question (e.g. *"Did you exercise today?"*), or some other form of input (e.g. *"Take a picture of your snack*

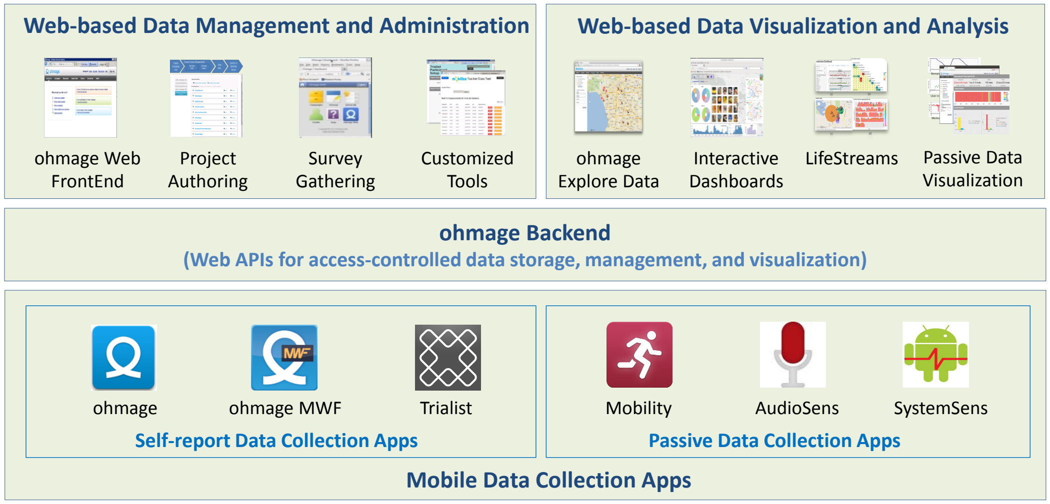


Fig. 1: Ohmage system architecture and its end-to-end PS software suite.

wrapper"). Ohmage supports a rich set of prompt types including: 1) single/multiple choices with the option for users to define additional choices; 2) number; 3) free-text; 4) timestamp; 5) multimedia such as pictures, videos or audios; and 6) a *remote activity* that launches a third party application, such as a game app to assess attention. Each prompt and message can optionally set a condition based on previous prompt responses that determines whether or not it should be displayed. For instance, a prompt *"Did you take your medication before, after, or without food?"* can be set to display if the user responded *"Yes"* to the previous prompt *"Did you take medication X this morning?"* This capability enables a more interactive and responsive survey answering experience.

A project can be scripted in the XML format that conforms to the ohmage project schema [Ohmage 2010a]. It can be manually created or automatically generated by an interactive Project Authoring tool (described in Section 3.4). Along with the XML content, there are also metadata (e.g. ID, name, creation time, running state, etc.) associated with a project. Once a project is validated and uploaded, it will become available for participants to collect data. Upon submission, all survey responses are validated against the project definition for correctness. Each uploaded survey response has an associated privacy state (i.e. private or shared); by default, it is set to private.

Passive data

Passive data refers to data streams that are passively and continuously collected from mobile device sensors or applications. An *observer* defines a set of passive data streams including their payload schemas and metadata. The payload schema describes the structure and data types of the payload. The metadata determine whether a unique ID, timestamp and location will be provided for each payload record. Similar to a project, an observer can be scripted in the XML format that conforms to the ohmage observer schema [Ohmage 2010a]. Once an observer is uploaded, passive data streams conforming to the observer definition can then be submitted to ohmage. All passive data are validated against the observer definition for correctness. There are currently three mobile apps that submit their passive data to ohmage: Mobility, AudioSens, and SystemSens (described in Section 3.3). We are also working on integrating third-party sensing apps, such as funf [Aharony et al. 2011a], to use the observer APIs.

3.2. Ohmage Backend

The ohmage backend is the central component of the ohmage platform. It provides secure communication, authentication, account management, access control, data storage and management, data analysis, and visualization through an extensive set of backward compatible APIs. We describe its key ohmage functionality below.

Account management and access controls

The ohmage account management and access controls are designed to support proper individual data access in concurrent diverse deployment settings. An ohmage account can be created by an ohmage administrator, or the user herself through a self-registration mechanism. Users are authenticated based on their account credentials. In addition to individual account management, ohmage provides a group management capability allowing a collection of accounts, each with a different group role (i.e. privileged vs. restricted), to be managed together. When a project is added to a group, the project and its data are automatically available to all the group members, each with different access privileges depending on their assigned project roles (i.e. supervisor, owner, analyst, participant). For example, participants are only allowed to submit survey responses, analysts can have read access to participants' shared data, and supervisors have full control over the data. The members' project roles are automatically derived from their group roles (e.g. all restricted members are given participant and analyst roles), but can be further adjusted to address specific settings. Beyond group- and project-based roles, ohmage also provides a number of global roles for extra operational privileges, such as a system administration, group creation, and project creation.

The group roles, project roles and user global roles are used by the ohmage access control mechanism to determine user access privileges in the system. Group management can be used to easily isolate sets of users and projects from one another, thus allowing multiple studies to be securely managed and deployed on the same backend. Finally, a combination of user roles and the ohmage data access control mechanism allows ohmage to flexibly address different deployment needs.

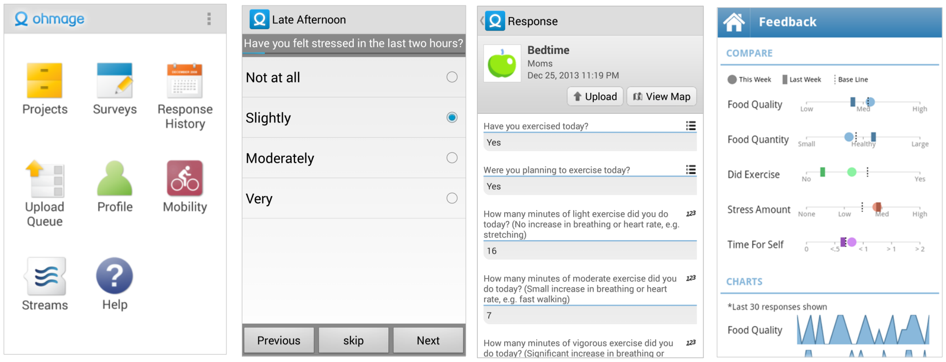
Data storage and management APIs

The main functionality of the ohmage backend is to provide create, read, update and delete (CRUD) operations to support the storage and management of all ohmage objects. To support consistent security across all applications, access to these CRUD operations by phone apps, administrative interfaces, and all other third-party tools is orchestrated solely through the ohmage web APIs [Ohmage 2010a]. As the backend evolves to provide more functionality, these APIs remain backwards compatible (within the same major release). Such unified and backward compatible interfaces enable the modularity and extensibility of ohmage by allowing different components to be independently developed and integrated without affecting other components.

Before any interaction occurs, an application must first authenticate its user with the ohmage backend and obtain a token for subsequent requests. Afterwards, all data access are subjected to the user's roles and ohmage access control mechanism. All ohmage APIs support HTTP compression (i.e. gzip format) for better speed and resource usage. For the read requests of self-report and passive data, ohmage also supports pagination and searching based on a time period to allow client applications to manage the returned data size. All transactions to the backend are logged in an audit log for deployment monitoring and analysis.

Data Analysis and Visualization

Along with the CRUD APIs, the ohmage backend also provides a set of APIs for primitive analytical and visualization functions of the collected data. These APIs are offered



(a) (b) (c) (d)

Fig. 2: Ohmage Android app: (a) The app home screen showing different menu options. (b) A prompt for user input. (c) A completed survey. (d) User feedback showing their current status versus their baseline in the Moms Study (Section 4.2).

through the backend subcomponent, OpenCPU (www.opencpu.org), which is a statistical computation software system, including analysis, visualization and reporting functionality based on R [Ooms 2011]. Client applications can access this set of APIs in the same way as they access other ohmage CRUD APIs. Examples of basic visualization functions are scatter plots of two prompts, prompt distributions, etc. These analysis and visualization APIs are extensively used in data visualization tools, such as the ohmage Explore Data.

Open-mHealth API compliance

Besides the interfaces defined in the ohmage platform, the ohmage backend also supports third-party APIs so that ohmage can be integrated into larger ecosystems. Specifically, the ohmage backend serving as a Data Storage Unit (DSU) implements Open-mHealth APIs which supports mHealth data interoperability and software reusability [Estrin and Sim 2010]. The APIs standardize data access so that client software can access any Open mHealth data store in a uniform way. Through these open APIs, open and closed source software components, as well as applications and data can be *mixed and matched*, allowing for meaningful insights to be derived through reusable data processing and visualization modules.

3.3. Mobile Data Collection Apps

The ohmage mobile data collection component consists of mobile apps that interact with the participants and collect various data streams. Based on the data streams collected, these apps can be classified into two categories: self-report and passive data collection apps. We describe each category of apps below.

3.3.1. Self-Report Data Collection Apps

Self-report apps allow participants to record their observations or experiences by answering a set of survey questions. A self-report app downloads the project definitions from the ohmage backend, and renders an appropriate graphical interface based on prompt order, type, and display condition. Survey responses are automatically timestamped, geocoded and uploaded to the ohmage backend. Three self-report apps, described below, have been developed on varying platforms for different scenarios.

Ohmage Android App

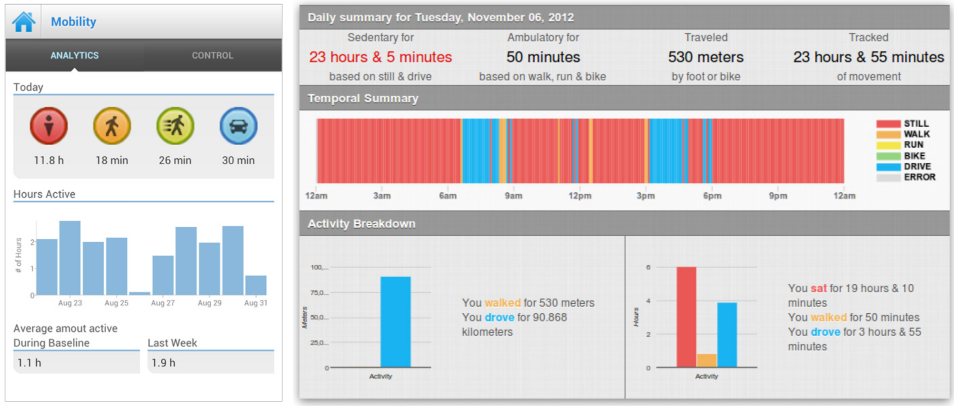
The ohmage Android app supports self-report data collection, temporally- and spatially-triggered reminders, and data summary; it also supports upload management of passively collected data. For self-report data collection, the app allows users to browse and participate in available projects in real-time. It fully supports all message and prompt specifications described in Section 2, including the video/audio capture and a third party app integration through the remote activity. For example, instead of describing one's emotion in text or rating, a project can utilize the Photographic Affect Meter (PAM) app [Pollak et al. 2011] that allows users to choose a picture, among a set of pictures, that reflects their current mood. The PAM input, along with other survey questions, are submitted to the ohmage backend as a survey response. Survey responses can be submitted even when the network connection is not available. These responses will be put in an upload queue. The app periodically checks the upload queue and uploads them in the background when the network is available. Figure 2a, 2b, and 2c show the app home screen, a prompted categorical question, and a completed survey response, respectively.

A few additional features are supported to facilitate the survey taking activity. Shortcut buttons can be configured for users to quickly submit short pre-filled surveys from the Android Home Screen. For example, a ”stress” button enables participants to report their stress with just one tap. The app also provides an interface for users to set reminders based on time and locations to take surveys. In addition, users can view their previous survey responses based on time (i.e. calendar view) and location (i.e. map view), and visualize a data summary that compares their current performance with a configurable baseline (shown in Figure 2d). This comparison is one instance of data feedback intended to sustain user engagement and participation; an area of active research in the ubicomp and mobile health communities.

Finally, the app also provides a unified upload manager to facilitate the upload of data collected via external passive data collection apps. The ohmage upload management includes secure upload, error handling and retry mechanisms. To minimize resource consumption, the manager can be configured to upload only via Wi-Fi and will not upload any data if the battery level is below 20%. All three ohmage passive data collection apps utilize the ohmage upload manager for their data submission to the ohmage backend. The ohmage Android app is available in the Google Play Store.

Ohmage MWF

One of the ohmage design goals is to allow participants to use their own existing smartphones, by supporting ohmage on multiple mobile platforms. To achieve this goal, we implemented a version of ohmage using the UCLA-developed Mobile Web Framework (MWF) [UCLA 2012]. Ohmage MWF is a lightweight cross-platform mobile web app that provides a subset of functionality of the ohmage Android app, namely self-report data collection and time-based reminders. It makes use of HTML5, javascript and MWF web technologies for dynamic and responsive page rendering. The HTML5 local storage is used to support offline operations (e.g. storing survey responses before they are uploaded). In addition, the PhoneGap library is used as middleware for the app to access native device features such as GPS, camera, and reminders across different platforms. Ohmage MWF is currently available on the Google Play Store, and Apple iTunes. The app will be extended to support richer multimedia prompts such as video and audio, spatial reminders, and visualization. As with most mobile web applications, ohmage MWF compromises look and feel for cross platform functionality. For this reason, we are also developing an iOS native ohmage app to be released mid 2014.



(a) (b)

Fig. 3: Mobility: (a) The app analytics screen showing mobility and ambulatory summaries for one and ten days, respectively. (b) The web-based Mobility Dashboard showing daily summary of user's mobility status, including the numbers of ambulatory minutes, distance on foot or bike, mobility times-series and activity breakdown based on time and distance.

Trialist

Trialist is a customized version of ohmage MWF that is designed to support n-of-1 randomized trials comparing two different treatments. It provides self-report data collection, daily, weekly and event-based reminders, and user feedback. The event-based reminders (e.g. when to start or switch between each treatment) are automatically set based on the randomized schedule configured via the Trialist Setup (Section 3.4). The user feedback shows the individual's experiment status (e.g. progress and time remaining). This app has completed pre-pilot testing and will launch in April 2014.

3.3.2. Passive Data Collection Apps

Passive data collection apps are mobile apps that collect data from sensors or other applications on mobile devices. Along with the raw data, apps may also provide high-level features (e.g. mobility mode, or voice versus non-voice) extracted by applying specific algorithms to the raw data. The main design criteria of these apps is to reliably collect data while consuming minimal resources, such as battery and network, to minimize the impact on the participants' phone availability. In this section three apps and their approaches for resource efficiency are presented. Each of these apps collects a different type of data and has a dedicated web-based tool (described in Section 3.5) to analyze and visualize the data.

Mobility

Mobility infers the user's mobility mode—namely: still, walking, running, and driving (or riding a bus)—using accelerometer, Wi-Fi, and GPS data. It uses the variance and FFT features of acceleration per second to classify whether the user is walking or running; if not, the app then uses Wi-Fi data to distinguish between driving and being still. If the phone detects that the nearby Wi-Fi ambient signals are static, the app assumes that the user is still; otherwise, it assumes driving. Wi-Fi data gives better accuracy and energy saving. However when the Wi-Fi signal is not available, GPS data is used instead. This classifier runs on the phone so the user can see the results in real-time. The classification results and the raw data are uploaded to the backend for further analysis and visualization. To further reduce the energy consumption, Mobility duty-cycles the accelerometer and GPS, if used. The user can choose to sample the mobility data every one or five minutes. With a five-minute sampling period, the phone can usually last throughout the whole day without being charged. A data control and visualization page is available for users to adjust the sampling rate and visualize its data summary. Figure 3 shows the app data summary and the web-based Mobility visualization.

AudioSens

AudioSens samples and processes audio data obtained from the microphone on the phone. It uses a speech classifier to detect whether an audio sample contains speech or not. This high-level feature can be used to investigate a user's social interactions [Lu et al. 2009]. For example, AudioSens was used to investigate the interactions among family members in a familial study. To reduce the energy consumption, the app dynamically adjusts the sampling rate according to the frequency of speech. When the app has not detected speech for over a certain amount of time (e.g. when the user is sleeping), it exponentially decreases the sampling rate. In addition, the app provides a data visualization page for users to review their data. All detected features are uploaded to the ohmage backend.

Systemsens

Systemsens captures and analyzes the phone usage data. It records over 30 types of phone usage information covering 1) event-based records (e.g. phone calls and messages) generated whenever a system state changes, and 2) polling-based records (e.g. CPU, memory, and battery usage) generated periodically by taking a snapshot of the system states. These data can be used for troubleshooting during the deployment or analyzing phone usage pattern [Falaki et al. 2011]. For example, it has been used to investigate the communication patterns among support group members of gay men. All the records are timestamped and stored in the JSON format and periodically uploaded to the ohmage backend.

3.4. Web-Based Data Management and Administration

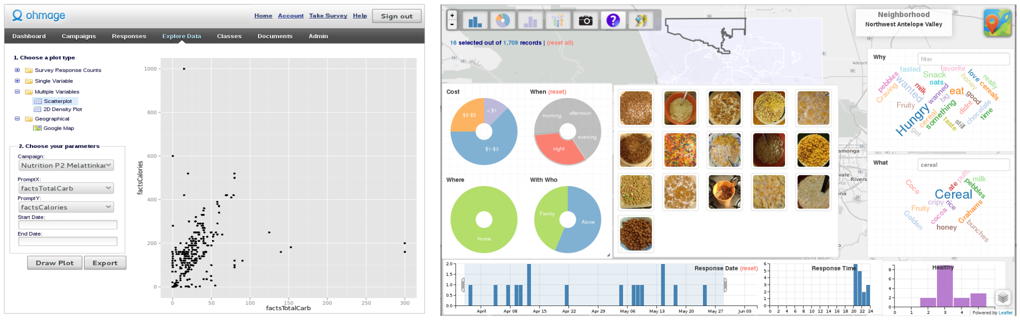
Multiple web applications have been built in the ohmage platform to support data management and administration. Most applications can be used in any deployment, with a few exceptions that are currently customized for specific deployment needs. These applications are described below.

Ohmage Web Frontend

Ohmage Web Frontend (or Frontend) is the main project, data, and user management portal implemented based on Google Web Toolkit. A system administrator can use the Frontend to create and manage user accounts and groups, as well as audit the ohmage log for anomalies. A project coordinator or manager can use the Frontend to create, modify, export, or delete projects. Once created, a project automatically becomes available for participants to join and submit responses. Participants can use the Frontend to view or delete their data, change privacy states of their responses, and export their data in the CSV format. A self-registration interface is available for new users to create their own accounts. The ohmage Web Frontend also has a unified login system whereby other web applications on the same server can simply redirect their logins to the Frontend login and allow it to handle all client authentication.

Project Authoring

A project definition can be scripted in an XML file. However, for non-technical users such as researchers or teachers who are not familiar with programming, a Project Authoring tool provides a graphical user interface. It guides the user through a step-by-



(a) (b)

Fig. 4: The ohmage Explore Data and public Snack Interactive Dashboard. (a) A scatter plot of the Snack Nutrition data collected by students in a Math classroom showing Total Carbohydrates versus Total Calories. (b) A dashboard showing multiple single-variable distributions (e.g. Snack Cost, Healthy Level, images, etc.) of the Snack data filtered by ”Cereal” snacks that users had at night.

step project creation process and automatically creates the project XML content that can be uploaded to the ohmage backend. Users can click to view the temporary XML content at any time during the process. On-screen tool-tips describe each configuration parameter as well as potential values for that parameter.

Survey Gathering

Even though PS focuses mainly on the use of mobile devices as data collection tools, in some deployments such as Mobilize (Section 4.1), not all participants have access to those devices. To allow such a group of users to participate in the data collection, a generic web-based application, extended based on the ohmage MWF codebase, was created. This tool allows users to complete their surveys using desktop browsers with two drawbacks: 1) there is no photo taking capability; instead, a user will be asked to upload a picture; and 2) the GPS obtained from the browser might be inaccurate or unavailable. Despite its limitations, this tool has proven to be very useful in the educational deployments.

Customized Setup

Some deployments require additional setup process that couldn't be satisfied with the existing ohmage management tools. Two interesting tools—Trialist Setup and Group Management Tool—have been developed for such deployments. Trialist Setup provides an interface for clinicians to configure individual n-of-1 randomized experimental trials of two treatments. A randomized experimental schedule is automatically created based on the entered configuration parameters, and is used by the Trialist mobile app to setup proper treatment reminders. Group Management Tool allows the delegation of group management from an ohmage administrator to project coordinators. A coordinator, once given proper privileges, can dynamically create and update groups using a CSV file containing user information. Section 4 describes the motivations and how they are used in actual deployments. Though the tools are currently customized, they can be generalized to support similar studies in the future.

3.5. Web-based Data Visualization and Analysis

One of the most challenging tasks of any data collection project is data analysis. Instead of a one-size-fit-all solution, ohmage provides a set of tools that facilitate data exploration, visualization and analysis for different usage scenarios. For self-report data, there are 1) generic ohmage Explore Data which is readily available for all projects, and 2) Interactive Dashboard which can be customized for specific projects. For passive data streams, there is a specific visualization for each ohmage passive data collection app. Finally, Lifestreams [Hsieh et al. 2013] takes both self-report and passive data into its data analysis software stack and tries to make sense of both types of the data simultaneously. All tools dynamically retrieved their data from the ohmage backend. This section describes each of these tools.

Ohmage Explore Data

Ohmage Explore Data allows project managers and end-users to explore and visualize high-level analytics of self-report data collected within an arbitrary project. There are three different views provided: 1) project monitoring view that focuses on the data collection progress, as well as user participation (e.g. leader board); 2) spatial map view displaying locations of survey responses associated with one or all users with zooming capability; and 3) prompt detail view that provides single variable time-series and distribution plots of different prompts associated with one or all users, and two-variable scatter and density plots of all users. An example of a two numerical scatter plot is shown in Figure 4a. The application supports limited subsetting functionality; only allows subsetting based on time and users. All the plots are created by the OpenCPU of the ohmage backend. Ohmage Explore Data is readily available as part of the ohmage Web Frontend.

Interactive Dashboards

The Interactive Dashboard was designed to address the ease-of-use and subsetting limitations of the ohmage Explore Data. It follows the *”Overview first, zoom and filter, then details-on-demand”* design pattern [Shneiderman 1996]. On the dashboard, multiple tailored graphs (i.e. pie charts, bar plots, maps, wordcloud, image pane) each representing a different survey question can be simultaneously displayed. Users can directly interact with individual graphs to filter and zoom into the data. Specifically, users can select a range within a bar plot, choose a segment of a pie chart, or type in a specific word in a wordcloud to filter the data. Multiple filtering conditions can be applied simultaneously to further zoom into the data. Once a graph is filtered, other graphs will be immediately updated accordingly. As a concrete example, Figure 4b shows an interactive dashboard customized for the Mobilize Snack project (Section 4.1) that can be used to explore the healthy-level rating of different snacks during different time of day. Interestingly, based on the data collected by students in Spring 2012, ”cereal” was rated healthier in the morning than at night.

The self-report data are dynamically retrieved from the backend. However, all the processing and rendering are done on the browsers. A JSON configuration file is used to set how variables in a project are displayed. This file can be customized and deployed for any project by a system administrator. A public version of the Snack Interactive Dashboard mentioned above is available at http://lausd.mobilizingcs.org/snackdemo.

Lifestreams

The ultimate goal of many PS studies is to make sense of the data and make the results available and actionable to the end users. Lifestreams [Hsieh et al. 2013] is an analysis stack that tries to achieve that goal. The Lifestreams data analysis stack consists of four layers including feature extraction, feature selection, inference and visualization. Feature extractions transform the raw ohmage data into more meaningful information, e.g. location traces can be transformed into places and duration-of-stay, etc. Feature selections allow researchers to select relevant features either manually or through the provided statistical methods (e.g. correlation analysis). The inference modules are used to identify behavioral trends and patterns based on the selected features. Specifically, the behavior change detection module identifies behavior changes and estimated change points. The correlation summary module identifies behavioral patterns through correlation between different features. Finally Lifestreams visualization allows researcher to explore participants' everyday life patterns by clicking through various plots. Lifestreams was used to identify behavioral patterns in a long-term behavioral study (Section 4.2) that analyzed self-report and passive data streams.

Passive Data Visualization

Ohmage provides three different visualization dashboards for the passive data streams collected by Mobility, AudioSens, and SystemSens apps. The Mobility dashboard displays a daily summary of participants' mobility status, including the numbers of sedentary and ambulatory minutes, distance on foot or bike, mobility times-series and activity breakdowns based on time and distance (shown in Figure 3b); mobility maps; and historical analysis that looks at activity breakdowns over time. The AudioSens dashboard shows individuals' speech patterns over time and space, comparisons of speech patterns between multiple users, and aggregated speech time statistics for a specified period. Finally, the SystemSens dashboard shows time-series of different resource consumption such as battery level, network, and CPU usage, and the user's interaction with the phone, such as screen on and off. These dashboards are used to provide feedback to participants. They are also used by the developers and deployment managers for application debugging and troubleshooting purposes.

4. OHMAGE CASE STUDIES: SUPPORTED RESEARCH AND FINDINGS

ohmage has played a significant role in a number of independent behavioral, health, wellness and education research projects. Examples include: high risk behaviors in HIV+ men, young adults with ADHD, exercise in south Asian women, mood and energy in breast cancer survivors, post-traumatic stress disorder (PTSD) in veterans, smoking cessation in teenagers, monitoring diabetes in a personalized randomized control trial (N-of-1), etc.

In this section, we present three distinct ohmage use cases: 1) Mobilize: Mobilizing for innovative computer science teaching and learning; 2) Moms: Studying diet, stress, and exercise related risk factors for cardiovascular disease in young mothers; and 3) PREEMPT: N-of-1 trials using mHealth in chronic pain. These use cases demonstrate different research disciplines, captured data, targeting populations, and scale. Mobilize presents a broader use of end-to-end participatory sensing practice in a larger educational deployment setting. The Moms study presents a typical application of ohmage in behavioral and wellness studies that collect and analyze participants' Ecological Momentary Assessment (EMA) and passive contextual data. Finally, PREEMPT demonstrates a role of ohmage in personalized, mobile-based, n-of-1 randomized experiments.

4.1. Mobilize: Mobilizing for Innovative Computer Science Teaching and Learning

The Mobilize project (www.mobilizingcs.org) is a targeted National Science Foundation Math Science Partnership project funded for 2010 to 2015. Mobilize aims to engage STEM (Science, Technology, Engineering and Mathematics) students in data collection and analysis activities that promote computational thinking and civic engagement via innovative math and science instructions. At the heart of Mobilize is Participatory Sensing, in which students use mobile phones and web services to systematically collect and interpret data about subjects important to them and their communities.

In Mobilize, students develop their computation thinking skills through the following PS process: 1) create their data collection projects by choosing the topic and designing the details of the projects; 2) collect data by making key decisions about what,

Table I: Mobilize deployment summary showing numbers of classes, students, and survey responses across different subjects and years.

Deployment Year	ECS (classes)	Maths (classes)	Science (classes)	IDS (classes)	Total (classes)	Total (students)	Total Survey Responses
2011-12 (completed)	13	-	-	-	13	377	3,821
2012-13 (completed)	12	7	5	-	24	644	6,250
2013-14 (on-going)	9	45 (44 to-date)	33	-	87 (44 to-date)	3,045 (1,242 to-date)	5,261(to-date)
2014-15 (expected)	10	90	70	10	180	6,300	-

where, and when to report observations; and 3) explore/analyze the collected data. Ohmage has been designed to equip LA Unified School District (LAUSD) STEM classrooms with an integrated and iterative learning environment for students to practice the PS process. Curriculum units in four different subjects have been developed or are under development to incorporate this learning process: 1) an 8-week Exploring Computer Science (ECS) Data Analysis unit involving research question formulation, project creation, data collection, and data analysis; 2) a 3-week Math unit on linear algebra content; a 3-week Science unit on biology content; and 4) a full-year Introduction to Data Science (IDS) course featuring lessons newly created to introduce students to data science and statistics.

Ohmage, along with the Mobilize curricula, have been deployed to a total of 81 LAUSD STEM classrooms, and 2,263 students since Spring 2012. Table I summarizes the numbers of classes, students and survey responses in different subjects and deployment periods. Estimated numbers of classes and students are provided for future deployments.

ohmage Application

For project creation, two canonical projects—Snack and Advertisement—were created for each ECS class in previous deployments. Starting from 2014, with the improved curriculum, students are expected to create their own projects by explicitly creating the XML definition files or using the ohmage Project Authoring tool. For the Math and Science classes, due to the limited instructional time, the Snack Nutrition and Trash projects are pre-created for those classes respectively.

For data collection, Mobilize relies primarily on the ohmage self-report data collection tools. During the Spring 2012 deployment, only the native Android app was available. To address the technology accessibility issue among students, 250 Android phones with data plan were loaned to the ECS students. This stopgap was expensive in terms of financial and managerial overheads and would not scale well in a larger deployment. Moreover, some students complained about carrying an extra phone (e.g. they often forgot to carry or charge the phones). Since then the ohmage software suite had been expanded to provide a cross-platform self-report mobile app and a browser-based survey gathering capability. In 2013, all students were asked to use their own mobile devices including smartphones, iPod Touch, iPad/tablets and school computers for their data collection exercises.

For data exploration, analysis and visualization, Mobilize relies on the ohmage Explore Data and the Interactive Dashboards. The Interactive Dashboards are easy to navigate and explore and can be used to quickly engage students in data inquiry. Alternatively, the ohmage Explore Data allows students to systematically explore different plots, including scatter plots of two numerical variables that Math students can use to observe linearity (Figure 4a). In addition, RStudio—an open source software that provides an interactive graphic user interface for programming and analyzing data in R—is also used in ECS and IDS classes for in-depth data manipulation and analysis.

An important, but often overlooked, step behind a successful deployment is the accurate and timely setup of user accounts and groups (or classes in this context). Typically, the ohmage Web Frontend is used by an administrator to centrally create and manage the setup. However, in a large deployment such as Mobilize where there are hundreds of classes and their members keep changing, the centralized approach is inefficient and doesn't scale well. A more efficient approach is to delegate the setup task to individual teachers, so they have control over their classes and simultaneously reduce the project administration overhead. Thus, ohmage has been extended to provide the Group Management tool for teachers with proper privileges to create their own groups (or classes) and dynamically manage their student members whenever their rosters change. The tool supports the CSV files exported from the LAUSD roster system and has been in place since September 2013.

Results and Findings

Based on the Mobilize project evaluation of the past deployment [Ong et al. 2012; Ong et al. 2013], the results have been encouraging. For the 2011-2012 implementation, the ECS teachers found the unit to be innovative, enjoyable, as well as technically and pedagogically challenging. Student also had positive learning experience with the unit. For the 2012-2013 implementation, the ECS teachers felt more confident in engaging students in deeper discussions related to data. Math and Science students found the units engaging and developed a better appreciation of the role of data in their lives.

In terms of technology, among 192 ECS students in 2012, most students found the technologies to be easy to use: from the scale of 1 to 5 where 1 is very hard to use and 5 is very easy to use, the average scores were 4.1 for the Android app and 3.8 for the ohmage Web Frontend. As for what they liked most about the unit, 15.1% liked everything about the unit, technology, and/or the new learning experiences; 28.7% students liked the unit due to the use of smartphones; 27.6% liked taking pictures and/or collecting data; 20.3% liked learning more on the project topics e.g. their snacking habit and health implication or advertisements around their communities; 16.1% liked seeing shared data, analyzed and/or made conclusions on their data. An example of what a student liked most about the unit is "I enjoyed the fact that we were able to personally collect various data about our own lives as well as recording them with phones, rather than pencil and paper."

The Mobilize project has successfully demonstrated a broader use of ohmage as an innovative teaching and learning software in STEM education. In addition to LAUSD high school classes, ohmage has also been piloted in a few undergraduate political science classes at UCLA in 2013, and is in the process of offering it as a UCLA-wide instructional tool.

4.2. Moms: Studying diet, stress, and exercise related risk factors for cardiovascular disease in young mothers

Cardiovascular disease (CVD) is the leading cause of death amongst women. One of the major risk factors for CVD is weight. More than 60% of women in the US are overweight [NIH 2013]. This situation is worsened when women become mothers, which typically increases BMI each year. Such trends emphasize the need for prevention or reduction in CVD risk factors for mothers beginning in young adulthood. Diet, exercise, and stress reduction can reduce CVD risk factors, but more affordable, valid, reliable, and feasible tools to self-monitor these behaviors are required to address the limitations in current methodology. The Moms study aims to assess the validity and reliability of using mobile technologies (i.e. ohmage) to monitor CVD risk factors in young mothers, and secondarily to evaluate the efficacy of using the smartphone for behavior change.

Through both self-report and continuous sensing apps, the study collected participants' diet, stress, mood, exercise and activity traces throughout the day. 56 young moms participated in and completed the pilot between January 2012 and March 2013. Among them, the experimental group consisted of 44 moms who used our apps to collect data for 6 months—the first 3 months of data were used to establish an individual baseline, and during the subsequent 3 months, moms received feedback that compared their current reports with their established baselines.

ohmage Application

During the 6-month pilot, each mom was given a smartphone installed with the ohmage Android app and Mobility app. Surveys were launched 4 times a day with reminders to prompt the participant to report their diet, stress, mood and exercise routines. Shortcut buttons to capture meals and stress events in real-time were available when moms did not necessarily have time to complete a survey, but wanted to record the time and location of an event. After the baseline period, participants would get a positive reinforcement feedback on behaviors that showed improvement compared to the established baselines. Participants could also go to the feedback screen to compare their current performance with last week's performance and their baselines (Figure 2d). In addition, moms were asked to run the Mobility app to collect continuous accelerometry, GPS, Wi-Fi signals, and the inferred mobility states. The ohmage Frontend was used by the researchers to create the project. The Frontend and the Explore Data visualization were used by the study coordinator and researchers to manage and monitor the on-going study. The Lifestreams visualization was later used for a qualitative study involving 8 out of 44 moms during in-person interviews.

Results and Findings

Mom's study is a relatively ambitious data collection project in terms of survey frequency as well as duration. 15,599 survey responses were collected across the 44 moms. The responses are distributed uniformly across the four surveys, 4,248 (27%) morning surveys, 3,968 (25%) midday surveys, 3,722 (24%) late afternoon surveys, and 3,661 (24%) bedtime surveys. On average, participants answered 2 surveys per day during periods where participants answered at least once. A total of 115,228 survey questions were answered, with an average of 2,619 questions per participant and 16 questions per participant per day. For contextual sensor data, 3,834 days worth of mobility data were collected with a minimum of 5, a maximum of 202, and an average of 87 days. This information on participation can be useful for other studies aiming for similar scale.

Almost all participants reported that the study made them more aware of their eating behaviors and conscious of unhealthy eating habits. The surveys and buttons as well as the positive feedback received at the end of surveys helped increase moms' awareness of their habits. As a result, some moms sought out professional help to change their behaviors. In addition, participants and the study coordinator found Lifestreams helpful in providing more insight into an individual's behavior data and in guiding the discussion with the participants. The behavior change detection algorithm was deemed as one of the most useful analysis. More detail of the algorithm and detailed findings are provided in [Hsieh et al. 2013].

There were a few common issues that arose during the deployment in asking participants to use the study phones. Participants reported problems with battery life, primarily due to forgetting to charge the phone since it was not their primary phone. The reminder to charge the phone after the last survey of the day was reported as being helpful. Additionally, some participants, those who didn't use the phone for other purposes outside the study, reported difficulty in remembering to carry their phone with them. This was especially a problem in analyzing moms' mobility traces. These lessons emphasize the needs for cross-platform mobile apps to support the use of participants' own phones in future studies.

The Moms study represents a general use case of ohmage application in health and wellness researches. Similar studies that used ohmage include the Family Wellness study that used the Android app and AudioSense to evaluate the feasibility of using mobile technologies for familial behavior studies; and a study on mobile phone-based assessment of health-related behaviors among gay men and their peers that used the ohmage MWF and SystemSens to monitor gay men behaviors and their social interactions (e.g. phone and SMS traces) among supporting groups [Comulada 2013]; etc.

4.3. PREEMPT (Personalized Research for Monitoring Pain Treatment): N-of-1 Trials Using mHealth in Chronic Pain

Chronic musculoskeletal pain is an enormous problem; affecting more than 100 million Americans [Boudreau et al. 2009]. Pain Treatments are often prescribed in a "trial and error" fashion which takes time to identify a successful treatment. N-of-1 trials are randomized controlled crossover trials conducted in a single patient. By crossing a patient back and forth between two treatments several times, clinicians can identify the more effective approach for that individual with greater precision than can be achieved in ordinary practice. PREEMPT seeks to take advantage of the "always on, always worn" relationship between people and their cellphones to allow Pain patients to participate in n-of-1 studies as part of their everyday life. Under PREEMPT, an application ("Trialist") will be created to provide patients and their healthcare providers to run personalized experiments comparing two pain treatments. PREEMPT will assess the effects of participating in a mobile n-of-1 trial versus usual care on clinical outcomes, participatory decision making, satisfaction, adherence, and health care cost.

ohmage Application

The Trialist is a software package consisting of three components: a backend, a Clinician-Facing component, and a Patient-Facing component. The ohmage platform is used as the foundation for the Trialist. The ohmage backend provides secure data storage and access. In addition, it has been extended to support the survey masking functionality allowing different subsets of surveys within the same project to be visible to different users. As a result, the Trialist can use the same project definition for all patients, but only show the efficacy measurement surveys that the individual patient is interested in.

For the clinician-Facing component, there are two web applications—the Trialist Setup and the Trialist Dashboard. The Trialist Setup application (Section 3.4) allows clinicians and patients to collaboratively design an n-of-1 pain treatment trial from a computer/tablet in a clinical setting. Clinicians can choose through drop-down menus the experiment configurations including treatments to be compared, medication schedule, duration of each treatment period (e.g. 1 or 2 weeks), number of treatment period pairs (e.g. up to 4 pairs), and n-of-1 efficacy measurements such as pain levels or side effects. The Setup application will automatically generate a randomized schedule and save it in the ohmage backend. At the n-of-1 trial's conclusion, the Trialist Dashboard displays the results in actionable items (e.g., the probability of long-term success with treatment A versus treatment B) so that the clinician and patient can make a joint and rational decision on the best treatment.

For the Patient-Facing component, the ohmage MWF app has been extended and customized for the Trialist mobile app to provide the following functionality: 1) allows patients to collect data on treatment, adherence and interim outcomes via configured surveys; 2) provides reminders informing patients of when to begin or switch a treatment, daily/weekly reminders for patients to enter clinical survey data, and periodic reminders about the time-left on the current treatment; and 3) provides feedback (e.g. study progress, time remains before completion).

Future Deployment

PREEMPT is currently under development. The system is expected to be ready for a live deployment in April 2014. About 250 patients will be recruited to participate in this study; half of the patients will use Trialist, and the other half will adopt the usual care. The study consists of n-of-1 trials conducted over 4-12 weeks, and 3-month, 6-month, and 12-month follow-ups.

The Trialist demonstrates how mobile health and ohmage can be used to support n-of-1 studies. A successful implementation of PREEMT will pave the way for broader use of mobile-based n-of-1 trials in clinical practice across other chronic health conditions. It will also promote more personalized patient-centered health care.

5. RELATED WORK

In this section, we discuss previous PS efforts, as well as self-report and passive data collection applications.

5.1. Participatory Sensing

Many PS efforts have successfully demonstrated the feasibility of applying the PS paradigm to scenarios where traditional approaches fall short because of time, cost, or complexity. These efforts can be divided into two categories based on user participation. The first category focuses on non-personal citizen science projects that engages the public or community members in the data collection activity. For example, What's Invasive [CENS 2006] is a PS app that allows national park visitors to report invasive plant species with geotagged photos. EarPhone is a participatory noise pollution mapping system for urban areas that use phones' microphones to collect audio samples [Rana et al. 2010]. Boyle Heights Project [Acker et al. 2010] is a PS project in which Boyle Heights residents used mobile phones to document environmental conditions and community assets.

The second category focuses on individualized data collected by users to address their personal issues related to health and wellness. For example, DietSense [Reddy et al. 2007] used snapshots automatically taken by the phone (hanging on the user's neck), along with a photo auditing system, to share diet-related photos (food, plate waste, etc.) with care providers. PEIR [Mun et al. 2009] combined personal location traces and geographic information to detect transportation mode and infer the users' impact and exposure to air pollution. Mappiness [MacKerron and Mourato 2013] studied how local environment affects happiness by using an iPhone app to collect over 3 million mood reports, GPS locations and ambient noise levels from over 45,000 people in the UK, while providing them feedback about their mood in different context.

While some PS systems are specifically designed for particular deployment scenarios, some are more general and can be applied in broader settings. Examples of these platforms are Ushahiti [Okolloh 2009], Open Data Kit [Hartung et al. 2010], and My-Experience [Froehlich et al. 2007]. Ushahiti was originally designed for Kenyans to report violence incidents primarily via SMS on feature phones. It has been extended to support similar reporting such as emergency responses through the use of simple surveys via SMS, email, twitter, and web-forms. Open Data Kit was designed to help semi-professional users (such as fieldworkers) in developing countries to collect scripted survey data using Android devices. Open Data Kit has been used in public health, environmental monitoring, and human rights abuse in many countries. MyExperience is a system for capturing scriptable self-report data and a specific set of passive data (i.e. phone usage, location traces, and calendar appointments) running on Windows CE platforms to study the mobile technology usage. MyExperience provides limited visualization capability.

Compared to these systems, ohmage supports a broader set of data allowing both scriptable survey and passive data to be collected by participants through the course of their everyday lives. Ohmage data collection tools are available for iOS and Android devices. In addition to self-report capability, ohmage apps also provide participation assisting features such as personal reminders and data feedback. There is also a rich set of data management, analysis and visualization tools for researchers and users to manage and explore their data. Finally, ohmage provides flexible account management and access control mechanisms to support diverse deployment settings.

5.2. Self-report capture with mobile phones

Self-report is a common approach in behavioral and health studies. Moreover, self-monitoring through self-report has been shown to encourage behavior changes and increase adherence by supporting self-awareness and self-efficacy [Larimer et al. 1999]. However, data captured by retrospective self-report are usually highly biased and inaccurate due to the limitations of autobiographical memory [Bradburn et al. 2000]. Mobile devices have enabled a structured, in-context, closer to real-time self-report, referred to as Ecological Momentary Assessment (EMA) [Shiffman et al. 2008] to monitor affect, cognitions, and behaviors in a person's natural environment. We are interested in the judicious use of self-report, in particular in the health domain, since significant health problems can be addressed in part through self-monitoring, such as diabetes and asthma [Karter et al. 2001; Pinnock et al. 2007]. Ohmage is designed to support these applications by providing the survey customization mechanism for researchers to capture self-report data in various forms and in real-time.

5.3. Passive data capture with mobile phones

With increasing sensing capability and pervasiveness, mobile phones have become a valuable tool for capturing passive data such as locations, activities, and proximity. Reality Mining [Eagle and Pentland 2006] gathered Bluetooth, GSM, and application usage data to build models for social systems. Ubifit uses passive activity traces to provide feedback encouraging participants to incorporate physical activity into everyday life [Consolvo et al. 2008]. Funf [Aharony et al. 2011b] is an extensible mobile data collection platform that captures various data signals, including sensors, location, ambient Wi-Fi, Bluetooth signals, and phone usage. It has been used to study social mechanisms in everyday life [Aharony et al. 2011b]. Ohmage provides a general framework (i.e. observer) to facilitate the collection and integration of any passive data streams. It also offers various passive data collection and analysis tools as described in Section 3.

6. CONCLUSION

A general, reusable participatory sensing platform can benefit many applications in multiple areas. It helps save time and resources and allows projects to concentrate on their primary objectives, instead of technology. In this paper, we have presented the primary system design criteria and architecture of ohmage, an open source platform that provides end-to-end PS functionality. We have also demonstrated the breadth and flexibility of ohmage in supporting different research disciplines through three distinct use cases in education and health.

Ohmage can be setup in multiple ways to support a PS deployment. A project can host its own ohmage instance by installing ohmage software components on a JVM supported operating system. Simple installation packages are available for Ubuntu, and Fedora [Ohmage 2010b]. Due to its flexible account management and access control mechanisms, ohmage can also be offered as a software as a service (SaaS) to projects that do not have resources or want to deal with system administration. For example, this service model is adopted by the Mobilize project to provide PS services for individual classes in LAUSD. In addition, there is currently a public sandbox of ohmage (play.ohmage.org) that allows individuals to experiment with their personal data collection. In this sandbox, users can create their own accounts, create new projects or participate in existing ones, as well as collect and visualize their own data.

Ohmage will continue to evolve to incorporate new technologies and functionality through its iterative design process. While a native iOS self-report app is being developed, the ohmage MWF will be extended to provide more features e.g. video and audio prompts, spatial reminders, and data visualization. In addition, we are investigating a scriptable personalized reminder mechanism triggered by the user's previous submissions and other contextual information. Using the Open mHealth architecture (openmhealth.org), we will integrate with other third party sensing apps such as funf [Aharony et al. 2011a] and Moves (moves-app.com). This integration will help broaden the sensing capability available for diverse studies. For data analysis and visualization, we are developing a web application that allows users to dynamically explore multidimensional relationships in survey data through the customizable composition of statistical plots. Through dynamic data exploration, users can gain more insight into their data. All these extensions will strengthen and broaden ohmage capability, which will in turn benefit a larger user community.

ACKNOWLEDGMENTS

This work was made possible with funding from UCLA Center for Embedded Networked Sensing (NSF #CCF-0120778), the National Science Foundation Math Science Partnership (NSF #DUE-0962919), the National Institutes of Health (NHLBI #5RC1HL099556), the Intel Science and Technology Center for Pervasive Computing (ISTC-PC), Cornell NYC Tech, and gifts from Google and Nokia. The authors would like to thank J. Burke and M. Hansen for their contributions to the concept of PS and its use in Mobilize; D. George, V. Hajdik, H. Falaki, M. Monib, T. Pham, K. Parameswaran, and E. Wang for their earlier software development contributions to the ohmage codebase; and R. Rocchio, E. Sakabu, and C. Trance for their support of the development of ohmage MWF and the ohmage quality control.

REFERENCES

Amelia Acker, Martin Lukac, and Deborah Estrin. 2010. Participatory Sensing for Community Data Campaigns: A case study. *Center for Embedded Network Sensing Tech Report* (2010).

N Aharony, A Gardner, and C Sumter. 2011a. Funf: an open sensing framework. (2011). http://funf.org/

N Aharony, W Pan, C Ip, I Khayal, and A Pentland. 2011b. Social fMRI: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing* (2011).

D Boudreau, Michael Von K, CM Rutter, and others. 2009. Trends in long-term opioid therapy for chronic non-cancer pain. *Pharmacoepidemiology and drug safety* 18, 12 (2009), 1166–1175.

NM Bradburn, AA Stone, JS Turkkan, CA Bachrach, and others. 2000. Temporal representation and event dating. *The science of self-report: Implications for research and practice* (2000), 49–61.

Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. 2006. Participatory sensing. In *WSW06 at SenSys 06*.

CENS. 2006. What's Invasive. (2006). http://www.whatsinvasive.org/

Connie Chen, David Haddad, Joshua Selsky, and others. 2012. Making sense of mobile health data: an open architecture to improve individual-and population-level health. *Journal of medical Internet research* 14, 4 (2012).

comScore. 2013. U.S. Smartphone Subscriber Market Share. (2013).

WS Comulada. 2013. Mobile phone-based assessment of health-related behaviors among gay men and their peers. *in submission to Connections, International Network for Social Network Analysis* (2013).

Sunny Consolvo, David W McDonald, Tammy Toscos, and others. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *CHI '08*. ACM.

Nathan Eagle and Alex Pentland. 2006. Reality mining: sensing complex social systems. *Personal and ubiquitous computing* (2006).

Deborah Estrin and Ida Sim. 2010. Open mHealth architecture: an engine for health care innovation. *Science* 330, 6005 (2010).

Hossein Falaki, Ratul Mahajan, and Deborah Estrin. 2011. Systemsens: a tool for monitoring usage in smartphone research deployments. In *MobiArch '11*. ACM.

Jon Froehlich, Mike Y Chen, Sunny Consolvo, Beverly Harrison, and James A Landay. 2007. MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In *MobiSys '07*. ACM.

Carl Hartung, Adam Lerer, Yaw Anokwa, Clint Tseng, Waylon Brunette, and Gaetano Borriello. 2010. Open data kit: Tools to build information services for developing regions. In *ICTD '10*. ACM.

Scott Heggen, Osarieme Omokaro, and Jamie Payton. 2012. Mad Science: Increasing engagement in STEM education through participatory sensing. In *Ubicomm '12*.

CK Hsieh, H Tangmunarunkit, F Alquaddoomi, and others. 2013. Lifestreams: a modular sense-making toolset for identifying important patterns from everyday life. In *SenSys '13*. ACM.

AJ Karter, LM Ackerson, and others. 2001. Self-monitoring of blood glucose levels and glycemic control: the Northern California Kaiser Permanente Diabetes registry. *The American journal of medicine* (2001).

Santosh Krishna, Suzanne Austin Boren, and E Andrew Balas. 2009. Healthcare via cell phones: a systematic review. *Telemedicine and e-Health* (2009).

ME Larimer, RS Palmer, and others. 1999. Relapse prevention: An overview of Marlatt's cognitive-behavioral model. *Alcohol research and Health* 23, 2 (1999), 151–160.

Hong Lu, Wei Pan, Nicholas D Lane, Tanzeem Choudhury, and Andrew T Campbell. 2009. SoundSense: scalable sound sensing for people-centric applications on mobile phones. In *MobiSys '09*. ACM.

George MacKerron and Susana Mourato. 2013. Happiness is greater in natural environments. *Global Environmental Change* (2013).

Mobilize. 2010. Mobilizing for Innovative Computer Science Teaching and Learning. (2010). http://www.mobilizingcs.org/

Min Mun, Sasank Reddy, Katie Shilton, and others. 2009. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In *MobiSys '09*. ACM.

NIH. 2013. Overweight and Obesity Statistics. (2013).

Ohmage. 2010a. ohmage API Documentation. (2010). https://github.com/ohmage/server/wiki/

Ohmage. 2010b. ohmage Installation Manual. (2010). http://ohmage.org/usage.html#developer_panel

Ory Okolloh. 2009. Ushahidi, or'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory learning and action* (2009).

C Ong, N Griffin, K Binning, G Delacruz, C Byrne, K Show, and E Redman. 2012. Mobilize project evaluation, fall 2012 report. *National Center for Research (CRESST)* (2012).

C Ong, N Griffin, M Hansen, A Madni, K Binning, C Byrne, E Redman, and N Shin. 2013. Mobilize project evaluation, fall 2013 report. *National Center for Research (CRESST)* (2013).

J Ooms. 2011. OpenCPU: Producing and Reproducing Results. *http://www.opencpu.org* (2011).

H Pinnock, R Slack, C Pagliari, D Price, and A Sheikh. 2007. Understanding the potential role of mobile phone-based monitoring on asthma self-management: qualitative study. *Clinical & Experimental Allergy* 37, 5 (2007), 794–802.

John P Pollak, Phil Adams, and Geri Gay. 2011. PAM: a photographic affect meter for frequent, in situ measurement of affect. In *CHI '11*. ACM.

Nithya Ramanathan, Dallas Swendeman, W Scott Comulada, and others. 2013. Identifying preferences for mobile health applications for self-monitoring and self-management: Focus group findings from HIV-positive persons and young mothers. *International journal of medical informatics* 82, 4 (2013), e38–e46.

RK Rana, CT Chou, SS Kanhere, N Bulusu, and W Hu. 2010. Ear-phone: an end-to-end participatory urban noise mapping system. In *IPSN '10*. ACM.

S Reddy, A Parker, J Hyman, J Burke, D Estrin, and M Hansen. 2007. Image browsing, processing, and clustering for participatory sensing: lessons from a DietSense prototype. In *Proceedings of the 4th workshop on Embedded networked sensors*. ACM.

Saul Shiffman, Arthur A Stone, and Michael R Hufford. 2008. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* 4 (2008), 1–32.

Ben Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*. IEEE.

Kate Starbird and Leysia Palen. 2011. Voluntweeters: Self-organizing by digital volunteers in times of crisis. In *CHI '11*. ACM.

UCLA. 2012. UCLA Mobile Web Framework. (2012). http://mwf.ucla.edu/