Evaluating Off-the-Shelf Technologies for Personal Health Monitoring: A Hands-On Workshop

Kay Connelly
Indiana University
connelly@indiana.edu

Kelly Caine
Indiana University
caine@indiana.edu

Katie A. Siek
University of Colorado
ksiek@colorado.edu

Julie A. Kientz
University of Washington
jkientz@uw.edu

Daniel O. Kutz
Indiana University
dokutz@indiana.edu

Rima Hanania
Indiana University
rhanania@indiana.edu

Danish U. Khan
University of Colorado
khand@colorado.edu

Eun Kyoung Choe
University of Washington
eunky@uw.edu

ABSTRACT
There is growing interest in ubiquitous tracking technology for personal health and wellness. Research in this area is vibrant and rapidly producing many innovations. However, with such growth there arises a need to evaluate the functionality and success of these innovations. The purpose of this workshop is to provide an opportunity for the UbiComp community to work together on a comparative evaluation of a wide spectrum of personal-health tracking technologies. The evaluation will yield recommendations for what metrics provide the most insights into the use, design, and functionality of features, data fidelity, research system integration, and user experience. Applicants will submit proposals to evaluate specific technology. If accepted, they will be provided with pre-defined and collectively proposed metrics to carry out those evaluations. The results from all evaluations will be discussed and compared at the workshop. We will broadly disseminate evaluation methods and outcomes to the greater pervasive health informatics communities.

Author Keywords Health, sensing, personal informatics, self-monitoring, ubiquitous computing

ACM Classification Keywords J.3 [Computer Applications]: Life and Medical Science – health. B.8.0 [Hardware]: Performance and Reliability – general. H.5.2 [Information Interfaces and Presentation]: User Interfaces – benchmarking, input devices and strategies, prototyping

General Terms Human Factors, Design, Measurement

INTRODUCTION
People are increasingly turning to ubiquitous tracking technologies and management applications to help them better understand their personal health. The quantified self movement [1]—people tracking personal health metrics such as exercise, stress, and sleep and sharing their experiences—is driven not only by patients, but also by lay-people who want to learn more about aspects of their health and wellness. Thus, anyone can be a *consumer* of quantified self-management technologies. Ubiquitous applications decentralize healthcare and wellness by empowering the consumer and shift the focus of care from treatment to prevention. Industry (e.g., patientslikeme.com, Fitbit, Microsoft HealthVault) and research communities ([2, 3]) support consumers by creating innovative platforms, applications, and visualizations to improve quantified self-management practices. New conferences (e.g., Quantified Self Conference, mHealth, Pervasive Healthcare, and IHI) and established conferences have created tracks to showcase research in this area and help establish a formal research community.

Although innovation is important in any field, a research community must also take time to reflect on the innovations to determine: how specific functionality should be measured; what technology is acceptable to researchers and end users; and what aspects of technology use produce positive health outcomes. Since research applications and prototypes are sometimes not stable enough to do a long-term deployment study and thorough evaluation, we believe that off-the-shelf or open-source technologies can provide researchers with an evaluation sandbox to determine what areas of design and development should be focused on during the development cycle. If an individual research group were to take on this evaluation process, it would take quite a bit of time and resources, thus we propose a workshop where the collective UbiComp community can work together to evaluate a subset of off-the-shelf technologies for personal health management.

The purpose of this workshop is to do a comparative evaluation of different technologies to make recommendations for what metrics provide the most insights into the use, design, and functionality of features, data fidelity, research system integration, and user experience. Before the workshop, workshop applicants will submit proposals to evaluate specific off-the-shelf technologies based on the metrics outlined here and other metrics that the researchers deem fit. Then, the workshop organizers will consider all of the submissions, group them appropriately, and share the final set of metrics that should be used by all workshop participants based on the technologies being evaluated. Finally, workshop participants will conduct the evaluation before the workshop so that the workshop time can be used to discuss metrics, findings, and outcomes. Through this process, the workshop will widely disseminate, via presen-

UbiComp'12, September 5-8, 2012, Pittsburgh, USA.
ACM 978-1-4503-1224-0/12/09.

tations and publications, the metrics used with the associated outcomes and provide the greater health informatics community with a thorough evaluation of current off-the-shelf technology to assist them with identifying what would meet their needs for their own studies.

CURRENT TECHNOLOGIES

The recent surge in personal health management has resulted in various ubiquitous personal health monitoring technologies. These technologies belong to different platforms including mobile, embedded, web-based, and standalone systems, and may be used to monitor various health metrics, e.g., physical activity, diet and nutrition, mood, sleeping habits, disease management, and health conditions. *Off-the-shelf technologies* implies any technology that is available to consumers for purchase from a commercial source that senses or monitors a physiological event. For this workshop, we restrict this definition to include those technologies that can be used for monitoring aspects relating to health or wellbeing and are flexible enough to be adapted for research purposes. Although we broadly categorize these technologies in the following classes, workshop applicants are encouraged to propose evaluation of technologies even if they are not present in the following list.

Due to their widespread and prevalent use in personal health monitoring, the first category of technology consists of (1) mobile phones that include smartphones and more basic phones. Smartphones have advance features such as digital/video cameras, GPS navigation, touchscreens, accelerometers, web browsers, and high-speed data access, while more basic phones are limited to simple communication features such as making and receiving calls and text messaging. The next category comprises of (2) Tablet PCs that are mobile computers, larger than mobile phones, and generally have an integrated touchscreen keyboard e.g., iPad and Samsung Galaxy Tablet.

After cell phones and tablet PCs, the third category includes (3) wearable and non-wearable health-monitoring devices that utilize various sensors to track vital signs and other physical activities. These sensors may include heart-rate sensors, blood pressure sensors, pulse oximeters, galvanic skin response sensors, temperature sensors, flex sensors, accelerometers, and gyroscopes. Examples of devices that utilize these sensors are Fitbit [4] and Q Sensor [5]. Fitbit tracks an individual's steps, calories burned, and sleep duration while Q Sensor measures emotional arousal via skin conductance, physical activity, and temperature.

Apart from mobile and sensory devices, commercial and academic sectors have developed different types of (4) websites to monitor and improve personal health. For the purpose of this workshop, we are most interested in tracking and social support websites such as YawnLog's sleep tracker [6], RunKeeper's exercise habit tracker [7], and DailyBurn's meal and exercise tracker [8]. We acknowledge there are other dedicated quantified self tracking or standard technologies that could be framed as such personal health tracking applications and technologies (e.g., personal health records; patient-oriented social support websites (e.g., www.patientslikeme.com); social networking websites (e.g., Facebook); development platforms that provide researchers tools to design and develop ubiquitous health monitoring system; location sensors (e.g., GPS systems); motion sensing devices (e.g., Kinect); and communication mechanisms (e.g., IrDA, Bluetooth, and Zigbee)), however to ensure that our scope is not too broad, we decided to not include these technologies in our initial evaluation calls.

The final list of technologies evaluated by the workshop will depend on the submissions received.

METRICS

We expect to collect several different types of information to assist researchers in evaluating off-the-shelf health monitoring technologies for their own projects. A complete list of evaluation metrics will be provided to workshop participants upon acceptance into the workshop, however the current list of metrics is available in the appendix. The specific metrics will depend on the class of technology. Currently, our metrics fit into four broad categories:

Functionality and Features

This category is primarily descriptive in nature, outlining both the intended and emergent purposes of the technology. It also includes a description of how the technology collects data (e.g., automated, user-initiated or some combination) and the nature of the data (e.g., data types and granularity).

Data Fidelity

An essential metric for any health monitoring technology is the accuracy of the collected data under different use conditions. To this end, an important contribution of this workshop is to identify and test different metrics and protocols to collect those metrics that could be used as benchmarks in the future.

Integration with Research Systems

One critical consideration when choosing a technology for a research project is how well the technology can be integrated into a research system. Metrics include items such as

- Customizability: how easy or difficult is it to customize the technology to a particular user group/environment?
- Development environment: what tools and developer's environment are available for the technology?
- Programming languages: what programming languages can be used to access the technology?
- Existing APIs: if the vendor provides APIs to access the technology or data, what functionality is supported through the APIs and how easy are they to use?
- Use of standards: does the technology make use of standards that make for easier integration with other technologies? For example, web standards or data exchange standards.

- Developer community: is there an active developer's community that can help when a researcher runs into problems?
- Troubleshooting: are there troubleshooting tools available to detect problems with the technology, or is the technology generally easy/hard to troubleshoot when there is a problem?
- Hardware reliability: how reliable is the hardware, especially under conditions that may be encountered in Ubicomp in-situ deployments?
- Technical support: does the vendor provide technical support? If so, how good is the support and how responsive is it? How much does it cost?
- Output data usefulness: how useful is the data that can be accessed by researchers?
- Visualizations and analysis tools: if the vendor provides visualizations or analysis tools of the data for the end user, how good are they?
- Scalability: how able is the technology to accommodate and handle growing amount of work?
- Stability: how likely is the technology to be stable for the foreseeable future (e.g., supported by the vendor, not outdated by software upgrades)?

User Experience – Consumer

Finally, the end-user experience is a crucial component when acquiring a technology for a research project. If the consumer does not like it or cannot use it, the research is likely to encounter difficulties. While the workshop timeline does not allow for user studies to answer these questions, it is possible for experts to perform initial evaluations. User experience metrics for the consumer include:

- Ease of use: how easy is the technology to use by the end user?
- Perceived usefulness: what is the user's perception about the usefulness of the technology?
- Aesthetics: if the technology is visible during normal use, how aesthetically pleasing is it?
- Responsiveness: how responsive is the technology to a change in what it is measuring?
- Form factor: is the form factor appropriate and likely to be acceptable by end users?
- User burden: how much user burden is there to operate and maintain the technology (e.g., charging batteries, synching data)?
- Number of steps for user to collect data: for the most basic uses, how many steps does it take for the user to collect data?
- Number of steps for user to access data: for the most basic uses, how many steps does it take for the user to access data?
- Privacy: is the technology designed with privacy in mind? Are there major privacy concerns that researchers should consider with the technology (e.g., a company that claims wide latitude is using user data, or a technology that displays sensitive data)?

While the organizers have developed this initial list of metrics, the workshop submissions will be encouraged to include additional metrics that can be incorporated into the evaluations.

WORKSHOP ORGANIZATION

Workshop participants will be expected to perform a detailed evaluation of one or more technologies before the workshop. Participants will be selected based on the technologies they propose to evaluate in their submitted paper and their ability to perform the required evaluation in the given time frame. Organizers will select participants to ensure a diverse coverage of technologies and will work with applicants to choose alternative technologies if multiple applicants propose to evaluate the same technology. The goal is to be as inclusive as possible and evaluate as many technologies as possible.

Once accepted, the organizers will provide workshop participants with specific instructions and standard evaluation metrics to ensure consistent reporting. At the workshop, sessions will be organized around specific technology themes, and data collected by workshop participants will be presented to the entire group.

The group will review the data and brainstorm additional data that may still be necessary to provide an accurate comparison between technologies within a theme. A final 2-hour session will have participants in specific themes break-out into groups to start preparing a publication based on the results. It is essential to have a substantial amount of time for this session so participants can make solid plans with respect to timelines, work contributions, and authorship. An organizer will lead each break-out group. Table 1 below lists key dates for the workshop.

May 18, 2012	Paper submissions due
May 25, 2012	Accept/reject notifications
June - July, 2012	Data collection
August 8, 2012	Evaluation data due
September 8, 2012	Workshop

Table 1: Key Dates

SUBMISSIONS

Each submission should be 2 pages or less and contain the following sections:

1. Technology Selection

Identify a high-level category of health monitoring technology that the authors would like to evaluate. If the technology is not obviously a health technology (e.g., location detection), the submission should include a justification for how it could be used within a health context (e.g., to support aging-in-place by tracking patients with dementia who may wander and get lost).

Please note that the workshop is not looking to evaluate *clinical* technologies. (e.g., EEGs, EHRs or clinical-decision support system) Instead, the focus is on technologies patients can use to monitor their health outside of the clinical domain.

The submission should indicate specific technologies the authors would like to evaluate and how flexible this technology selection is. The workshop does *not* have funding to provide equipment to participants for their evaluation, therefore the submission must indicate that the authors either already have access to certain equipment, or have a budget to acquire equipment. Since multiple submissions may request to evaluate the same technology, some workshop participants may be assigned to a different product within the same technology category. This will only occur with consultation of the authors before final acceptance. For this reason, it is important to indicate if the authors have already purchased the technology, and/or if the authors have a budget to acquire additional technology. Include the total number of units that are available for the evaluation.

2. Proposed Evaluation

The submission should specify the resources that the authors have available and the scope of the evaluation they believe is feasible with those resources (e.g., number of researcher hours and number of researchers who could use the technology during the evaluation).

The submission should also indicate the metrics that the authors would like to use during their evaluation. Include any metrics that may be unusual and/or unique to the research team (e.g., access to a physical fitness lab that can monitor respiratory status and metabolic rate).

In addition, if there are any classes of metrics the authors *cannot* use (e.g., perhaps they do not have someone qualified to perform a usability analysis), the authors should indicate that here. The workshop organizers will attempt to match different submissions so a complete evaluation of each technology can occur.

3. Author Information

The submission should include a brief bio of each author and indicate how they will contribute to the proposed evaluation. Ideally, author teams will be comprised of different expertise in order to accomplish the different types of evaluation that must occur.

REFERENCES

1. Swan, M., *Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking* International Journal of Environmental Research and Public Health 2009. **6**: p. 492-525.
2. Estrin, D. and I. Sim, *Open mHealth Architecture: An Engine for Health Care Innovation.* Science Magazine, 2010. **330**(6005): p. 759-760.
3. Medynskiy, Y. and E. Mynatt, *Salud!: An Open Infrastructure for Developing and Deploying Health Self-management Applications*, in *Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare (Pervasive Health 2010)*2010: Munich, Germany. p. 1-8.
4. http://www.fitbit.com/
5. http://www.affectiva.com/q-sensor/
6. http://beta.yawnlog.com/
7. http://runkeeper.com/
8. https://dailyburn.com/