On Usability Analytics and Beyond with Human-Centered Data Science

Marc Ericson C. Santos
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
chavez-s@is.naist.jp

Goshiro Yamamoto
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
goshiro@is.naist.jp

Jarkko Polvi
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
jarkko-p@is.naist.jp

Christian Sandor
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
sandor@is.naist.jp

Takafumi Taketomi
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
takafumi-t@is.naist.jp

Hirokazu Kato
Nara Institute of Science and Technology
Ikoma, Nara 630-0192, Japan
kato@is.naist.jp

Abstract
Extracting meaning from large volumes of data can possibly help HCI and CSCW researchers answer research questions better. In this position paper, we point to some of the challenges in conducting usability evaluations and how human-centered data science could help. Going beyond usability evaluations, we explain some paradigm shifts in HCI and CSCW that could benefit from human-centered data science.

Author Keywords
usability; evaluation method; handheld devices

ACM Classification Keywords
H.5.1. Multimedia Information Systems – Artificial, augmented, and virtual realities; Evaluation/methodology

On Usability Evaluations of Novel Systems
The central concept of the field of human-computer interaction or HCI is usability and usefulness [1]. To help us make sense of the usability and the usefulness of our systems, we adapt techniques from other fields like human factors, ergonomics, psychology, systems engineering, and computer science. Through the years, we have witnessed debates [2], [3] and proposals [4], [5] on how we could improve our practice – more

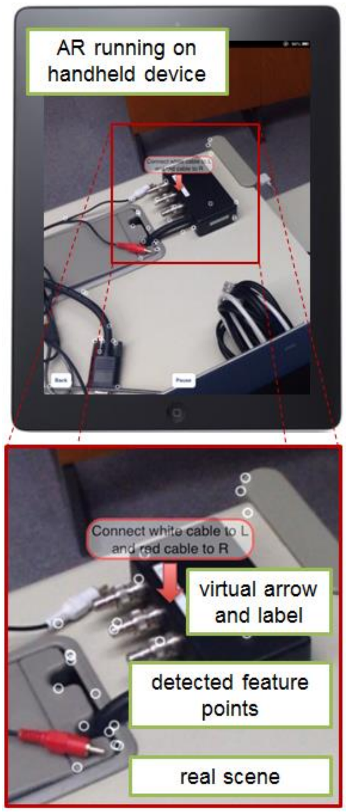

Figure 1: The HAR system we tested in [8]. The system runs on a tablet PC and enables the users to add virtual labels onto the real scene. More information about this HAR system, and our other HAR interactions can be found in [9].

specifically with regard to how we conduct our evaluations.

Applying inappropriate usability evaluation methods could suggest the rejection of an otherwise good design direction. Greenberg and Buxton [4] argue that traditional usability evaluation should not be used for validating early designs or culturally-sensitive systems. Instead, other reflective and critical methods could be applied. Compared to HCI, more research works in CSCW apply reflective and critical methods, such as the use of ethnography and field studies.

The application of human-centered data science in HCI might be more obvious for empirically validating designs and interfaces. That is, after implementing a prototype, we can draw insight from lots of quantitative data to support why our proposed interface is superior to its alternatives. The less obvious challenge for human-centered data science is how it could help researchers during the brainstorming phase of design and during the early system design. For example, how can we use human-centered data science to help us generate system requirements? How can we use data science to verify design guidelines to inform beginners and non-experts?

Drawing insight from lots of data is not a new idea. Websites are fine-tuned through web analytics – the measurement, collection, analysis and reporting of data related to the use of a website. In touch interfaces, Henze et al. have shown using millions of touch events that touch positions on a touch screen are systematically skewed [6]. We believe that this trend drawing insights from lots of data will continue not just for web clicks and touch events. For future HCI and CSCW research, it is important for us to study a variety of possibly meaningful data to improve usability analytics – the measurement, collection, analysis and reporting of data for the purpose of understanding and improving the usability and usefulness of a system.

In our ongoing work on handheld augmented reality (HAR), we argue that manipulability – the ease of handling the device – is an important factor in the overall usability of a system [7]. In other words, movement logs may be more meaningful in HAR than in other conventional uses of handheld devices. In a preliminary experiment, we demonstrate how we can estimate the user's usability rating of a HAR application based only on the accelerometer log of the system [8].

Our HAR application (Figure 1) requires the user to move a tablet PC from side-to-side to register SLAM feature points. The user then attaches a label on one of the points. This makes handling the HAR device challenging, which could reflect on the device's movement, more specifically on the device's accelerometer and gyroscope logs. After performing the task, we asked the users to answer a questionnaire. Our on-going work in [8] explores the use of accelerometer logs in creating decision tree models that could predict the usability rating assigned by users. We are interested to know if we can predict the user's perceived ease of use by observing how they move the device when using a HAR system.

Note that we do not advocate replacing behavior observations and the use of questionnaires. Insights generated from sensor log data should be compared with other information, such as insights generated from behavior observation, questionnaires, etc.

Beyond Usability

Researchers spend a significant amount of time to execute rigorous experiments that validate their design. Sometimes, these evaluations are motivated by passing the strict reviews of top conferences like CHI and CSCW. However, Chilana et. al. [10] argue that having usable and useful systems do not necessarily lead to adoption. In response, they offer a paradigm shift from a user-centered design to an adoption-centered design. They then document their lessons learned from transforming their research work to a product.

Assuming that we agree with adoption-centered design, data science will not only be used to evaluate usability and usefulness. It will also be used to prove business value for various stakeholders of a company or a community. Business value could be increased sales, reduced costs, extracted insights on consumer behavior, etc. Data science has been doing this for websites and web interfaces. How do we apply data science for other types of systems to substantiate business value? In the first place, should HCI and CSCW researchers concern themselves with adoption-centered data science?

Beyond individual users and well-defined collaboration groups, Lee and Paine [11] argue the need for new frameworks to capture the types of collaboration we deal with today. They suggest a paradigm shift with their Model of Coordinated Action (MoCA). MoCA extends Johansen's 1988 time-space matrix to include other aspects of CSCW, namely scale, number of communities of practice, nascence, planned permanence, and turnover. From this new model, they argue a shift from traditional notions of cooperation to coordinated action. The term coordinated action contains the meaning of people working together towards a shared goal. However, in coordinated action, the shared goal can be diffused and/or not defined clearly.

Lee and Paine [11] discussed Humanity Road to explain the new kind computer-supported collaborative work we have today. Humanity Road is a virtual organization of volunteers for humanitarian relief. Their work includes collecting and disseminating disaster information. This group has both episodic and long-term members who use various social media and collaborative tools, such as Skype, Twitter and Google Docs. The challenge for human-centered data science is how to capture coordinated action if:

- the participants are widely distributed and has a fast turnover rate
- individuals are contributing episodically without a well-defined goal
- the collaboration is happening beyond traditional groupwares
- the standard operating procedures of the community are still developing

Summary

Human-centered data science can help researchers improve their usability evaluation techniques. Beyond usability, we can consider how to conduct data science for adoption-centered design, and for capturing coordinated action or new types of collaboration.

In response to the theme of developing a research agenda for human-centered data science, we offer the following questions. How do we use data science…

- not only to validate our prototypes, but also in other design activities, such as during the gathering of system requirements?
- not only developing usable and useful systems, but also pushing for the adoption of our systems?
- to capture emerging types of collaboration?

Acknowledgements

This work is supported by the Grant-in-Aid for Japan Society for the Promotion of Science Fellows, Grant Number 15J10186.

References

1. H. Rex Hartson. 1998. Human-computer interaction: Interdisciplinary roots and trends. *Journal of Systems and Software*. 43, 2 (1998), 103-118.
2. Shumin Zhai. 2003. Evaluation is the worst form of HCI research except all those other forms that have been tried, essay published at CHI Place. Retrieved December 1, 2015 from http://www.shuminzhai.com/papers/EvaluationDemocracy.htm
3. Henry Lieberman. 2003. The Tyranny of Evaluation, essay published at CHI Fringe. Retrieved December 2, 2015 from http://web.media.mit.edu/~lieber/Misc/Tyranny-Evaluation.html
4. Saul Greenberg and Bill Buxton. 2008. Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08), 111-120.
5. Kasper Hornbæk, Søren S. Sander, Javier Andrés Bargas-Avila, and Jakob Grue Simonsen. 2014. Is once enough?: on the extent and content of replications in human-computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 3523-3532.
6. Niels Henze, Enrico Rukzio, and Susanne Boll. 2011. 100,000,000 taps: analysis and improvement of touch performance in the large. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI '11), 133-142.
7. Marc Ericson C. Santos, Jarkko Polvi, Takafumi Taketomi, Goshiro Yamamoto, Christian Sandor, and Hirokazu Kato. 2015. Toward standard usability questionnaires for handheld augmented reality. *IEEE Computer Graphics and Applications*, 35, 5 (September 2015). 66-75.
8. Marc Ericson C. Santos, Takafumi Taketomi, Goshiro Yamamoto, Gudrun Klinker, and Hirokazu Kato. 2015. Towards estimating usability ratings of handheld augmented reality using accelerometer data. In *Proceedings of the International Symposium on Mixed and Augmented Reality* (ISMAR '15). 196-197.
9. Jarkko Polvi, Takafumi Taketomi, Goshiro Yamamoto, Arindam Dey, Christian Sandor, and Hirokazu Kato. 2015. SlidAR: A 3D positioning method for SLAM-based handheld augmented reality. *Computers & Graphics*. 55 (2015). 33-43.
10. Parmit K. Chilana, Andrew J. Ko, and Jacob Wobbrock. 2015. From user-centered to adoption-centered design: a case study of an HCI research innovation becoming a product. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '15). 1749-1758.
11. Charlotte P. Lee and Drew Paine. 2015. From the matrix to a model of coordinated action (MoCA): a conceptual framework of and for CSCW. In *Proceedings of the SIGCHI Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '15). 179-194.