UNDERSTANDING & EVALUATING UX OUTCOMES AT SCALE

Serena M. Hillman

Microsoft, serena.hilman@microsoft.com

Samira Jain

Microsoft, samira.jain@microsoft.com

Craig M. MacDonald

School of Information, Pratt Institute, cmacdona@pratt.edu

Elizabeth F. Churchill

Google, churchill@acm.org

Carolyn Pang

Oracle, carolyn.pang@oracle.com

Jofish Kaye

Carelon Digital Platforms, acm@jofish.com

Erick Oduor

Maji.co.ke, erick@oduor.net

Evaluating the success of applications from a summative perspective is essential to many industry researchers' roles, yet a thorough understanding of what to measure and how to bridge business frameworks remains elusive. New technologies and novel ways of interacting with applications have garnered domain-specific interest in evaluating these experiences, but there is yet to be clarity or a path forward for a general methodological approach. While experience outcomes, UX OKRs, UX metrics, UX health, etc., are essential elements in HCI research and design, it has not become a core topic in HCI courses. This workshop aims to address this gap by bringing together academics and industry researchers involved in the measurement of success around large-scale applications – applications with a large user base. The objective of this workshop will be to engage in dynamic discussions around current evaluative experiences, how we move forward to produce valuable and unified updates to this topic, and a plan to do so. The workshop will specifically look at the value of writing a State of the Union paper, developing a community (e.g., SIG), and exploring education and training opportunities around this topic.

CCS CONCEPTS • Human-centered computing • Human computer interaction • HCI design and evaluation methods

Additional Keywords and Phrases: Summative research, Experience outcomes, KPIs, UX metrics, UX health, OKRs, UX outcomes, user-centered metrics

1 BACKGROUND

Measuring the success of large-scale applications from a summative perspective is a critical part of many HCI/UX research teams' objectives. Large-scale applications offer the ability to collect data from a large user base and thus enable the collection of large volumes of behavioral data. The collection of such data is important as it can be leveraged in different and unique ways that smaller usability sample sizes cannot. When discussing UX metrics, experience outcomes, UX OKRs and general UX health, we are referring to the broadest possible view of UX. This includes social constructs like trust, credibility, value, diversity, and inclusion, etc., in addition to more traditional constructions like satisfaction, usability, engagement, ease of use, and usefulness. In our perspective, these metrics could be described as UX, user-centered or HCI-focused with the importance being that they measure the human experience and not business outcomes.

This topic presents significant and prevalent challenges that encompass various aspects. These challenges include, but are not limited to, selecting metrics that genuinely measure user experience rather than business outcomes (referred to as experience outcomes), determining appropriate data capture mechanisms for different metrics and user interactions, establishing the frequency of data collection, ensuring secure storage and governance of data in compliance with privacy regulations like GDPR, leveraging data to drive product improvements, addressing social and cultural obstacles to achieve buy-in for well-defined experience outcomes, tracking meaningful data over time to assess impact, maintaining consistency in sampling and data capture across different products within the same organization and across different organizations.

Furthermore, additional challenges often arise when seeking organizational support, particularly when organizations prioritize financial health indicators such as revenue and subscriptions over human-centric outcomes like enhancing users' experiences. To address these challenges, researchers must develop a comprehensive methodological approach to measure the impact on users and effectively communicate the long-term value and success of user experience achieved.

Over the last several years, this complex problem space has had little advancement within HCI. Instead, there has been a trend to focus on specific contexts and domains (e.g., mixed reality [3], disabled children [4], aging populations [7], and surgery [6]). This leaves general foundational works, situated almost ten years ago, as the primary contributions to this topic (e.g., [1][2][5]). Further, these foundational works focus primarily on selecting metrics and methods for collecting data. They do not discuss several challenges mentioned above (e.g., socializing, aligning with business metrics, and current organization culture). There is one exception to this trend, a case study by Hillman et al. [9] published in CHI 2022. This case study is authored by two of the workshop organizers and discusses the process employed and the lessons learned while designing and implementing a user-centered in-product feedback system within a family of large-scale applications. The paper identified challenges and opportunities around aligning with business outcomes, navigating current frameworks, unlocking self-serve data [8] to stakeholders, informing strategy, and feeding additional research [9]. We see this workshop as an extension to the Hillman et al. [9] case study, where we will explore others' academic and industry lived experiences of mapping experiences to metrics, reflecting on the nuances and complexities, and the knowledge gained.

Based on the identified knowledge gap and drawing inspiration from a previous case study, it is evident that this particular topic is primed for a concentrated and targeted endeavor to generate fresh insights and establish a clear direction. The proposed workshop aims to unite practitioners and academics, fostering discussions on the present state,

challenges, and accomplishments around this topic. By bringing together this diverse group, the workshop will facilitate dynamic exchanges that elucidate existing challenges and uncover tangible opportunities in academia and industry. Ultimately, these interactions will yield practical outcomes that can be implemented effectively to drive progress and offer tangible benefits.

2 BENEFITS TO THE COMMUNITY

This workshop offers several benefits to the CSCW community. Primarily, we see an opportunity for healthy and future-focused conversations between academic and practitioner communities around this topic that needs strong alignment and forward movement. Secondarily, we see strong potential to explore new ways of overcoming the challenges mentioned above by coming together as a community rather than trying to solve them in silos. Further, we recognize that the value of new methods and approaches to evaluation is determined not only by the strength of their theoretical underpinnings but also by their practical applicability.

3 WORKSHOP GOALS

We intend to bring academic and industry researchers together reflecting on the workshop participants' diversity of perspectives and challenges around experiences, such as engaging with business frameworks like OKRs, as well as establishing and socializing UX metrics. We also have a goal to create an ongoing community. This would involve evaluating the need for a Special Interest Group (SIG) for ongoing collaboration and contribution to the topic, publishing findings from the workshop, and looking to establish education and training opportunities through HCI courses. To foster a community on this topic, we will ask participants to provide preferred methods of communication (e.g., email/Slack/social media). Our workshop goals are purposely broad to reflect the widespread challenges faced by evaluators and to attract a wider group of practitioners and researchers.

4 PARTICIPATION, INCLUSIVITY & ACCESSIBILITY

As we prioritize diversity and inclusion within our workshop, we will design a strong hybrid experience. This will include efforts to both passively collect applications for participation (e.g., via the CSCW 2023 workshop webpage, CSCW 2023 social media) and proactively seek out representation by directly contacting potential participants. We aim to include both experts and those less-experienced researchers with an interest in evolving the topic. The former may not have current experience in the topic but will be screened for high enthusiasm for shaping its future direction and a strong indication of applying the findings. Thus, potential participants may include academics, industry researchers, professors, artists, graduate students, data scientists, developers, project managers, etc. During the workshop, the virtual audience will be supported by one of the organizers dedicated to facilitating strong communication between the environments.

For accessibility, we will be following the McGill guidelines [10] around accessible workshop design. For example, we will ask participants for feedback on accessibility both before and after the workshop, provide a clear agenda and goals in the pre-package, schedule breaks every 2 hours, offer a hybrid workshop design, and include closed captioning.

We believe the above efforts will allow us to bring strong representation from academic and industry professionals and to provide an inclusive, safe space for sharing and action. To attend the workshop, participants will be required to complete all pre-activities detailed below.

5   PRE-WORKSHOP & WEBSITE

To participate in the workshop, a 2–4-page position paper will be required. After acceptance, participants will be asked to review all selected position papers before the workshop and complete an activity sheet to help jumpstart workshop activities during the workshop. These three requirements are listed below with additional details and deadlines:

1. July 6th: Call for participation is live on our workshop website http://uxoutco.me
   The website will present the workshop's goals, structure, important dates, and overall expectations. We will ensure all important information is not represented in images, and we will follow color contrast guidelines.

2. Participant position papers
   To be accepted into the workshop, interested participants will be asked to submit a 2–4-page position paper. This paper should reflect on their involvement in the topic, past experiments and learnings, the successes they have achieved, the challenges they have faced, and their objectives or plans for future initiatives. While the following prompts can provide guidance for writing the position papers, there is no obligation to answer them:
   - *What has been your experience measuring and evaluating the summative success of large-scale applications?*
   - *What was your involvement in the process? Who else was involved, and how?*
   - *What challenges have you faced? If you did address them, how did you address them?*
   - *How would you describe the journey you have undertaken to identify the metrics you use to evaluate the UX of your products or services?*
   - *How would you describe the journey you have undertaken to socialize your UX health metrics and/or the process of implementing them with business metrics?*
   - *Have you explored providing self-serve data to stakeholders? If yes, how did you do it? If not, why haven't you?*
   - *If you completed the process, what has transpired after implementation?*
   - *If you don't have experience with large-scale measurements for evaluating overall success, what is your interest and motivation in this field?*

   While we recognize that the term large-scale (or at scale) can be subjective, we encourage participants to determine if different applications apply based on the varying domains and industries and the volume required to measure UX outcomes. We also look forward to exploring the large-scale definition within the workshop and encourage the participation of all definitions of large-scale.

   Submissions will be reviewed with a focus on quality, originality, and potential to contribute to the workshop goals. The submissions will be reviewed by the Workshop Organizers as well as the Program Committee (see list of PC members below).

3. Pre-workshop package
   Before the workshop takes place, participants will receive a pre-workshop package. This package aims to gather participants' objectives for the workshop, collect information about any accessibility needs they may have, and provide them with the workshop agenda.

6   OUTCOMES & CONCLUSIONS

While the exact outcomes and conclusions will be dependent on details collected during the workshop, there are several potential outcomes we expect and will design the conversations to drive towards. As mentioned above, we intend to evaluate the need for the formation of an Experience Outcomes Special Interest Group, to focus on the broader idea of

what experiences mean and how we evaluate them, as well as to establish clear communication lines in developing a community of practice. Fostering this newly formed community will be our primary outcome.

Second, we will identify potential authors to publish the findings from the workshop in a "State of the Union" paper. This paper will focus on describing the current challenges, success, and discussion around future research the community has identified as the next steps.

Finally, we see an opportunity to explore education and training around this topic as both a workshop outcome and a long-term outcome of these efforts. For example, how do we train industry researchers who need to understand collecting and activating data around these topics? How can students learn how to understand and analyze these complex datasets in the context of a classroom setting?

In summary, our workshop has five concrete outcomes:

1. Identify potential authors to publish the findings from the workshop as a State of the Union Paper.
2. Evaluate the need for the formation of an Experience Outcomes SIG.
3. Establish clear communication lines for the development of this community of practice.
4. Identify interested community members for long-term interest in the topic.
5. Explore education and training opportunities around this topic.

Sharing these goals, objectives, and schedules will also be within the pre-workshop package, so participants have clarity on the workshop's purpose.

7 ORGANIZERS

Workshop organizers have helped architect this proposal, plan the workshop, and will help run the workshop either remotely or on-site at CSCW.

Dr. Serena Hillman (she/her) is currently a Senior Research Manager at Microsoft, where she supports a team of researchers who help drive the future of data and analytics services and applications such as Power BI, and Azure's data warehousing and analytics products. Her research has spanned analytics applications, video games, social computing, e-commerce, and educational technologies – all of which involved large-scale applications and the measurement of experience outcomes.

Samira Jain (she/her) is currently a researcher with Microsoft's Data Cloud Studio, focusing on large-scale data analytics services and tools. Before moving to Seattle, she lived in seven different cities across the U.S. and India, consulting on various design and research challenges across various industries. For the past three years, she has been helping drive the implementation of a large-scale in-product feedback system across Azure Data products and services.

Dr. Craig MacDonald (he/him) is an Associate Professor in the School of Information at Pratt Institute, where he developed the Master of Science in Information Experience Design program and directs the Center for Digital Experiences, a student-driven, faculty-led UX consultancy, and academic research lab. His research focuses on building organizational UX capacity in practical settings, strengthening HCI/UX education, and building a community of practice for HCI educators.

Dr. Elizabeth Churchill is a Director of UX at Google. With a background in psychology, Artificial Intelligence, and Cognitive Science, she draws on social, computer, engineering, and data sciences to create innovative end-user applications and

services. She has built UX teams at Google, eBay, Yahoo, PARC, and FujiXerox. Her current focus is on the design of effective developer tooling.

Dr. Carolyn Pang (she/her) is a Senior Researcher with Oracle and brings over 15 years of experience working with enterprise resource planning (ERP) systems, including Oracle and SAP products. Her research has spanned social and urban computing, with a focus on building experiences that foster connection and a sense of community amongst large user bases in digital cities.

Dr. Jofish Kaye is Senior Director of User Research & Interaction Design at Carelon Digital Platforms, part of Elevance Health, formerly Anthem.ai. He runs teams that focus on building better tools to empower doctors, clinicians, and patients to manage their healthcare, and draw insights from large health datasets.

Dr. Erick Oduor is currently the Founder at https://www.maji.co.ke/, an App connecting water vendors and buyers in Kenya on an easy-to-use digital platform, cultivating a trusted and efficient community. Ex-IBM HCI and UX researcher focusing on ICTD. He was also VP UX at Research Strategy group in Toronto, Canada.

8 PROGRAM COMMITTEE

In addition to the main group of workshop organizers, we have also engaged a group of experts to help shape the workshop program as members of the program committee (PC). The PC includes representatives from both academia and industry and each member brings deep expertise or experience in one or more areas relevant to our workshop topic. In addition to helping the organizers build the final workshop program, PC members will also be involved in promoting the workshop as they see fit (e.g., via their professional networks and soliciting submissions).

Dr. Carman Neustaedter, Dean of the Faculty of Communication, Art, and Technology (FCAT)
Dr. Barry Brown, Research Professor at the University of Stockholm
Dr. Mark Schlager, Staff UX Researcher on the Google Fitbit UX team
Angela Moulden, Director of UX Research for Google Assistant
Dr. Kostas Kazakos, UXR Lead for Identity Verification in Google Payments Platform
Dr. Jessica Tran, Staff Human Factors Engineer at Apple
Devesh Desai, Director of Design Research at Oracle
Eric Liu, Staff UX Researcher at Meta
Ali Rizvi, Data and UX practitioner, PhD student at the University of Waterloo

9 REFERENCES

[1] Avinash Kaushik. 2013. See-Think-Care-Do: A Content, Marketing, Measurement Business Framework. (July 2013). Retrieved June 30, 2023 from https://www.kaushik.net/avinash/see-think-do-content-marketing-measurement-business-framework/

[2] Caitlin Sadowski, Margaret-Anne Storey, and Robert Feldt. 2019. A Software Development Productivity Framework. In: Sadowski C., Zimmermann T. (eds) Rethinking Productivity in Software Engineering. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4221-6_5

[3] Dmitry Alexandrovsky, Susanne Putze, Valentin Schwind, Elisa D. Mekler, Jan David Smeddinck, Denise Kahl, Antonio Krüger, and Rainer Malaka. 2021. Evaluating User Experiences in Mixed Reality. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 87, 1–5. https://doi.org/10.1145/3411763.3441337

[4]   Emeline Brulé, Oussama Metatla, Katta Spiel, Ahmed Kharrufa, and Charlotte Robinson. 2019. Evaluating Technologies with and for Disabled Children. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19). Association for Computing Machinery, New York, NY, USA, Paper SIG08, 1–6. https://doi.org/10.1145/32900607.3311757

[5]   Kerry Rodden, Hilary Hutchinson, and Xin Fu. 2010. Measuring the user experience on a large scale: user-centered metrics for web applications. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). Association for Computing Machinery, New York, NY, USA, 2395–2398. https://doi.org/10.1145/1753326.1753687

[6]   Roman Bednarik, Ann Blandford, Feng Feng, Antti Huotarinen, Matti Iso-Mustajärvi, Ahreum Lee, Federico Nicolosi, Jeremy Opie, Soojeong Yoo, and Bin Zheng. 2022. Integration of Human Factors in Surgery: Interdisciplinary Collaboration in Design, Development, and Evaluation of Surgical Technologies. In Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 112, 1–7. https://doi.org/10.1145/3491101.3503709

[7]   Sayan Sarcar, Cosmin Munteanu, Neil Charness, Jussi Jokinen, Xiangshi Ren, and Emma Nicol. 2021. Designing Interactions for the Ageing Populations – Addressing Global Challenges. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 110, 1–4. https://doi.org/10.1145/3411763.3441326

[8]   Sean Byrnes. 2019. What is Self-Service Data Science? Forbes, Innovation, Retrieved June 30, 2023 from https://www.forbes.com/sites/forbestechcouncil/2019/08/22/what-is-self-service-data-science/

[9]   Serena Hillman, Samira Jain, Vichita Jienjitlert, and Paula Bach. 2022. The BLUE Framework: Designing User-Centered In-Product Feedback for Large Scale Applications. In Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 21, 1–8. https://doi.org/10.1145/3491101.3503558

[10]   Student Services McGill University, 8 Tips for Making your Workshop more Accessible, Retrieved from https://www.mcgill.ca/skills21/files/skills21/accessibility_workshop_pdf.pdf