The Role of Design in Creating Machine-Learning-Enhanced User Experience

Conference Paper · March 2017

1 author:

Qian Yang
Carnegie Mellon University

The AAAI 2017 Spring Symposium on
Designing the User Experience of Machine Learning Systems
Technical Report SS-17-04

The Role of Design in Creating Machine-Learning-Enhanced User Experience

Qian Yang
Human-Computer Interaction Institute, Carnegie Mellon University
qyang1@cs.cmu.edu

Abstract

Machine learning (ML) applications *that directly interface with everyday users* are now increasingly pervasive and powerful. However, user experience (UX) practitioners are lagging behind in leveraging this increasing common technology. ML is not yet a standard part of UX design practice, in either design patterns, prototyping tools, or education. This paper is a reflection on my experience designing ML-mediated UX. I illustrate the role UX practice can play in making machine intelligence usable and valuable for everyday users: it can help identify 1) how to choose the right ML applications. 2) how to design the ML right. The separation of the two concerns is a first step to untangling the tight interplay between ML and UX. I highlight the unique challenges and the implications for future research directions.

Introduction

Machine learning (ML) applications *that directly interface with everyday users* are now increasingly diverse and pervasive. Many of the most popular mobile apps regularly mine user behavior and contextual data in order to make personalized recommendations, filter out spam, predict travel time, and log behaviors like walking or sleeping. The maturity and prevalence of this technology catalyzed the notion that *ML is the new UX* -- that is, ML will be the most important way to improve user experience (Brownlee, 2015; Kuang, 2013; Yang, Zimmerman, Steinfeld, & Tomasic, 2016).

Designers are lagging behind in leveraging this not-so-new technology. The shape of the ML-mediated future seems to have been mostly driven by data availability and learner performance rather than a deliberate user-centered vision. ML has not yet become a standard part of UX design practice, nor a part of design patterns, prototyping tools, or design education. UX designers do not have enough tacit understanding of ML to operationalize interaction flows that proactively adapt or evolve over time (Yang, Zimmerman, Steinfeld, & Tomasic, 2016).

How should UX design practice utilize and instantiate ML? How can UX practitioners purposefully use machine intelligence to fulfill user needs and adequately weave them into real-world contexts? How does design deal with the UX uncertainty and temporality associated with ML algorithms?

These fundamental UX questions underlie my research in HCI. Over the past 3 years as a researcher at Carnegie Mellon University, I have concentrated on designing intelligent systems for real world uses. In particular, my research involves 1) designing ML-driven software that helps clinicians with patient selection for artificial heart implant surgeries (Yang, Zimmerman, & Steinfeld, 2015; Yang, Zimmerman, Steinfeld, Carey, & Antaki, 2016); 2) adding intelligent UI adaptations to a mobile app for users with disabilities (Yang, Zimmerman, Steinfeld, & Tomasic, 2016), and 3) designing an autonomous driving UX (Basu, Yang, Hungerman, Singhal, & Dragan, 2017). These systems interact with end users directly in high-consequence real-world scenarios, where the user's perception and experience of ML are vital to the systems' success.

In this position paper, I reflect on my experience designing these systems, then map out the roles UX can play in making machine intelligence usable and valuable *in the wild* for *everyday users*. I highlight the unique challenges that designing ML-mediated UX entails, as well as the implications for future research pathways.

Machine Learning as a Design Material

Case Study: Designing Clinical Decision Support

One of my ongoing research projects is designing an ML system that aids high-risk clinical decision-making, also known as *decision support tools* (DSTs). The system mines thousands of patient records, bringing the collective intelligence of many physicians to each implant decision of whether and when to implant an artificial heart into an end-stage heart failure patient (Yang et al., 2015; Yang, Zimmerman, Steinfeld, Carey, et al., 2016). Previously, such systems have taken a context-less, prototypic form: A list of patient condition measures is fed in and produces an individualized prediction of patient trajectories, such as likely survival and other post-surgical risks (Bellazzi & Zupan, 2008). Interestingly, almost all these systems have failed when moving into clinical practice, despite their effectiveness in labs (Yang et al., 2015).

Given the previous failures of DST deployments and the wide gap between the prototypic technology and clinical reality, we chose to conduct a field study. We interviewed and observed clinicians caring for candidate patients at three different implant centers for 13 days. Our field study identified several barriers that could negatively impact the use and perceived value of machine intelligence, such as an attitudinal barrier. Physicians perceived no need for data support because they felt that they already knew how to effectively factor patient conditions into clinical decisions. They also lack trust in the ability of machine intelligence; instead, they rely on professional networks for actionable suggestions.

These observations in the field forced us to confront two fundamental questions in designing DSTs. Here I borrow the vocabulary from Schön and DeSanctis (1986):

1) What is the right thing to build? (Problem-setting). Take the attitudinal barriers as an example. In our study, physicians reported that they know how to make a clinical decision. Many implied that makers of current data-driven prognostic systems want to replace their expertise with inhuman technology. Taking a lesson from early HCI work in participatory design, we need to make technical advances that skill workers instead of de-skilling them. Under this problem framing, it becomes clear that, in order to motivate the clinicians to use the DSTs, we should design machine intelligence to skill clinicians rather than de-skilling them, to make them feel better with their decision making, rather than making the decisions for them;
2) How do we build this? (Problem-solving) The machine intelligence available does not always align with the need of users. In our study, we observed a mismatch between clinicians' information needs to make a good implant decision and DST's information flow. At the input end, prognostic models take in quantitative and explicit inputs, while challenging decisions are often characterized by unavailable or ambiguous medical and/or social evidence. Clinicians might find the information that most concerns them is not captured in the prediction, such as the patient's home life and social support, which are critical and difficult factors most often not captured in the medical history. In terms of DST data output, physicians need support for action taking. Consultation between cardiologist and surgeon best captures this: Is this case too risky to operate on? No? Ok, then do it. A probabilistic ML prediction can be obscure in telling whether to execute a therapy or not, or whether to do it now or to "wait and see".

Design Process

The separation of problem setting and problem solving is essential to carrying out a design process. Through an active process of ideating, iterating, and critiquing potential solutions, designers continually reframe the problem as they attempt to *make the right thing*. The final output of this activity is a concrete problem framing and articulation of the *preferred state*, and a series of artifacts—models, prototypes, products, and documentation of the design process (Zimmerman, Forlizzi, & Evenson, 2007).

One of the main reasons why ML is often difficult to design with is that there is no common way to separate problem setting and problem solving in a machine learning process. In general, the intrinsic qualities of available data anchor the capabilities and limitations of ML outputs, leaving little space for exploring design possibilities outside of the current paradigm.

Below I propose two fundamental dimensions along which to begin a discussion on the roles UX should play when designing with ML: learning the right thing and making the learning right.

UX's Role in Learning the Right Thing

> "To be design-oriented (in HCI) is to consciously seek to intervene and manipulate, aiming to convert an undesired situation into a desirable one." (Fallman, 2003)

UX Goals Converses with ML Goals

The articulation of "the desired future" is at the center of UX design. Early investment in activities like user studies helps designers ideate on divergent framing and solutions, and establish *what is the "right thing" to design*.

In contrast, machine learning efforts emphasize *what can be (accurately)* learned given the available datasets for a designated application. The shape of the ML-mediated future seems have been mostly driven by data availability and learner performance rather than a deliberate vision. The above case study exemplifies this inherent mismatch. Consequently, many existing ML systems fail to account for the users, or to fit into their context at large (Amershi, Cakmak, Knox, & Kulesza, 2015; Yang, Zimmerman, Steinfeld, Carey, et al., 2016).

Matching the machine learning capability and the right UX problem is essential and challenging. There seems a lack of dialogue between UX and ML expertise in system design processes as well as in ML development and training. It remains unclear what design processes or teamwork patterns can enable and facilitate such dialogues.

The first step towards such dialogue is to plan and seed the use of ML early in the system development process, *with goals explicitly articulated and thoughtfully seeded* (Yang, Zimmerman, Steinfeld, & Tomasic, 2016). Below I list several common lenses through which ML goals can be defined. This list is by no means meant to be complete; it is meant to be useful to dissect the consequences of machine intelligenc on UX.

- Through *users'* lens. For example, our design of the decision support tool has the explicit goal of helping clinicians feel they are doing better work, and not necessarily automating the part of work that makes them feel like an expert.
- Through *service providers'* lens. For example, recommender systems often aim for longer use sessions and more user engagement.
- Through a *humanistic* lens. This perspective of design goals addresses ML consequences at large. For example, how do media content rankers minimize the filter bubble effect (Nguyen, Hui, Harper, Terveen, & Konstan, 2014)? How can designers prevent "scoring" algorithms — systems that score teachers and students, sort resumes, grant or deny loans — from manifesting pernicious feedback loops that enforce biases? (O'Neil, 2016; Rader & Gray, 2015)

ML systems in the wild often need to cover and balance two or more of these perspectives. For example, Facebook's News Feed ranking algorithm has (at least) two design goals. One goal is to "show everyone the right content at the right time so they don't miss the stories that are important to them." The second is that the News Feed should display posts more prominently that will generate more interaction or "engagement" (Rader & Gray, 2015). Through users' lens, these two goals separate consumption (don't miss important stories) from production (increased engagement). Holistically, the two goals balance users' needs (information and social engagement) and Facebooks' need (longer user dwelling time and more user engagement).

Design Research Identifies New ML Opportunities

Design in early phrases *is* exploration. UX professionals often invest time on divergent work, essentially looking around in a design space of possibilities. They routinely consider user needs and desires, working to embody the people they make technologies for (Zimmerman, Forlizzi, et al., 2007).

Therefore, we argue that UX design approaches have great potential in identifying new ML opportunities. Designers can situate ML algorithms in different contexts and for different audiences. For example, in the case of designing DSTs, our field study revealed what the physicians' real information needs were when facing difficult patient cases. One example is the decision support for Emergency Room patient cases where data input is often sparse. We thus started mining relevant datasets that characterize this subgroup of patients and designing decision supports dedicated to urgent patient cases.

With adequate ML literacy, designers can potentially invent new ways of utilizing data, envisioning and operationalizing ML methods. While often times designers seem to lack such literacy (Dove, Halskov, Forlizzi, & Zimmerman, 2017), we have seen some exciting concepts emerging. For example, Leahu (Leahu, 2016) proposes a new, relational perspective in which HCI builds, tests, and engages neural networks in the design of interactive systems. Leahu proposes a hybrid approach where machine learning algorithms are used to identify objects as well as connections between them; finally, he argues for remaining open to ontological surprises in machine learning as they may enable the crafting of different relations with and through technologies.

More work needs to be done to enable UX practitioners to leverage ML as a design material. Currently, Practice-focused design patterns as well as UI/UX books and articles do not discuss how to design for ML adaptation or how to collect the information needed to do ML (Yang, Zimmerman, Steinfeld, & Tomasic, 2016). ML UX education and toolkits must provide new means that help UX teams develop a tacit understanding of ML -- that is, not simply teaching how ML works, but empowering them with enough technical literacy to be able to *ideate creatively* yet *practically,* and to better collaborate with machine learning experts.

UX's Role in Getting the Learning Right

UX Manifests the ML Pipeline

ML-mediated interfaces simultaneously shape user behaviors and drive its underlying machine learning pipelines, raising exciting new challenges and opportunities for designers.

When dealing with these interfaces, UX teams should consider what aspects of user behavior and context to capture in order to accurately externalize and interpret the interaction traces logged. They also need to consider how to

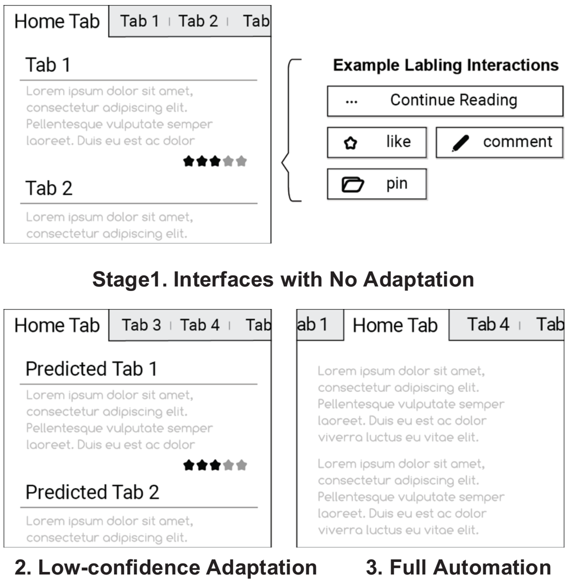

Figure 1. An Example of Adaptive UI Design Patterns - Landing Tab. On this set of interfaces, the data traces of user interactions can serve as a ready source for machine learning. (Yang, Zimmerman, Steinfeld, & Tomasic, 2016)

detect inference errors and even capture labels to enable reinforcement learning. Ideally, *UX professionals construct interfaces in such a way that the traces of user interactions can serve as a ready source for machine learning.*

Our recent work explores ways to achieve this ML-UI symbiosis in designing adaptive user interfaces. We provide a set of UI patterns (Figure 1 is one example of the pattern), that can 1) collect user traces needed to adapt or automate interface navigation; and 2) evolve as the use history data accumulates and as the confidence of ML prediction increases (Yang, Zimmerman, Steinfeld, & Tomasic, 2016).

This vision of ML-UI symbiosis marks an exciting space for future design and research in HCI. Our work narrowly investigates solutions in design patterns and prototyping tools. However, they represent only one small way that ML can enhance UX. HCI researchers have addressed this challenge by developing the ML system and interfaces simultaneously (Zimmerman, Tomasic, et al., 2007), allowing end-users to pick, modify and annotate training data (Amershi, 2011, 2012), etc. UX design practitioners also have been creating entertaining interactions that motivate users to contribute information that can be used as labels for personalization (Wood, 2014).

UX Choreographs User and System Evolvement

Creating interaction flows with two dynamic parts, the users and the machine interfaces, entail many new design opportunities and challenges. Specifically, ML enables interface representations and behaviors to change:

- Across different contexts,
- For different users,
- Over time.

ML-enabled adaptations can boost UX with agility and granularities that would otherwise be difficult to achieve. At the same time, they pose new challenges to UX sketching, prototyping and testing.

Appropriating UX for ML Performance

Recently, HCI researchers have started introducing human factors in to ML process. Kay, Patel, and Kientz (2015) raised the question of "how good is an 85%-accuracy classifier?", addressing the subjectivity of ML performances. The answer to this question comes down to the tradeoff between the quality of the learning system and the risk/reward users associate with a successful or failed prediction.

To design interactions that are appropriate into ML performances involved, designers should consider the potential benefits and costs of UX at each stage of the interface evolvement, and accordingly set an estimated accuracy

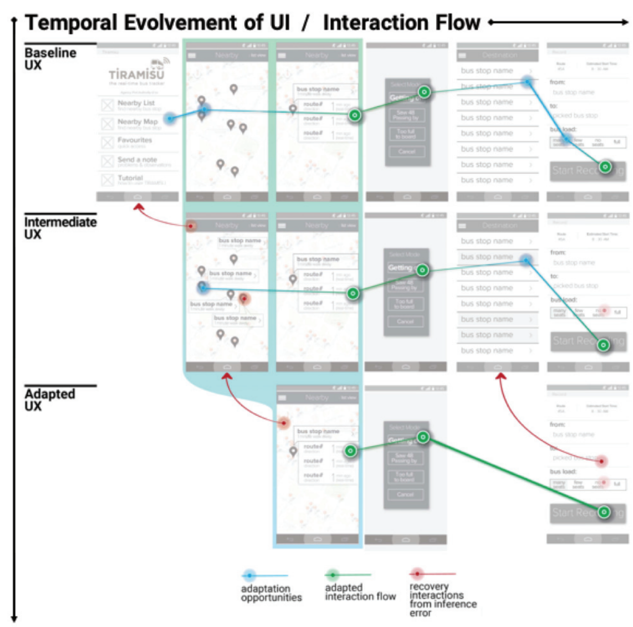

Figure 2. An Example of Adaptive UI Wireframes. Data collected over the early lifecycle of the app enabled great opportunities for ML to simplify in-app navigation. The blue dots represent user actions that designers hope to automate; the green traces are the user actions needed when all adaptations are successful; red traces represent page transactions to recover from an inference error. (modified from Yang, Zimmerman, Steinfeld, & Tomasic, 2016)

threshold required for triggering the next level of personalization or automation. Figure 2 is a preliminary example of how new prototyping tools can create a manifestation for these new design considerations. More work needs to be done to understand how users perceive the value of a successful adaptation as well as how to measure it against users' frustration with an inference error, under different use contexts and for difference types of errors.

UX Evolvement Over Time

Among the three factors to which intelligent UX adapts (context, user, time), I particularly highlight the design opportunities around the time dimension. ML's ability to evolve UX over time opens up exciting design opportunities for interaction and service designers. Data-driven personalization can build and maintain long-term customer relationships with a view toward creating lifetime value.

Interestingly, promises of future intelligent interactions often come at the expense of usability at the moment. In order to make confident inferences, ML systems often need more explicit input from users. Facing a new user, how can the interaction design induce more interactions, in order to collect high-quality ML features and labels? Facing a user with rich use history and context data, how can a designer find the harmony between automation and user engagement?

In prior work, HCI researchers have provided many insights on how to design slow technologies. However, there still is a lack of systematic understanding that is pertinent and apposite to the design of ML-mediated UX evolvement. I suggest that constructive design researchers can address this challenge by developing sensitizing concepts and strong exemplars that reveal new ways designers can unleash the opportunity of intelligent UX evolvement. HCI researchers can also put their rich body work in interactive machine learning tools in a temporal context, and reveal how people -- from developers (Patel, 2010), artists (Fiebrink, Cook, & Trueman, 2011), and makers, to end users in lab settings (Amershi, 2011, 2012) -- interact with UX that evolves over time.

Discussion

This paper elaborated on two fundamental roles UX practice plays when designing ML-driven systems: problem setting and problem solving. That is, to purposefully use ML to solve the right, user-centered problem, and to design the UI/UX in a way that produces and communicates ML outputs appropriately. To do so, designers should emphasize balancing the back-end concerns on both ML performance and usability (getting the thing right) with an up-front investment in exploration and ideation (getting the right thing).

Like others in the HCI and UX communities, I promote the idea that ML is the new UX. I envision UX practitioners leveraging machine learning (ML) as a design material creatively and thoughtfully, guiding users and technologists toward a deliberative ML-mediated future. Embedded in this vision is the inherent interdependencies between the ML expertise and the UX thinking: UX needs the context awareness and personalizability enabled by ML; ML needs UX design to be perceivably valuable to the users. Ideally, ML and UX exist in a symbiotic relationship with one another.

Toward this vision, more work needs to be done to advance current UX practices. I will continue to work on growing an understanding of and supporting the conversation between ML and UX teams. With these two fields working in concert, design goals can become grounded situations for future ML work, while ML methods can be appropriately applied to UX problems. I encourage the UX and HCI research community to join us in investigating new ways to evaluate the UX benefits and risk of ML and developing new ML capable sketching and prototyping tools.

References

Amershi, S. (2011). Designing for effective end-user interaction with machine learning. In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology - UIST '11 Adjunct* (p. 47). New York, New York, USA: ACM Press. http://doi.org/10.1145/2046396.2046416

Amershi, S. (2012). Designing for Effective End-User Interaction with Machine Learning. *Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*, *3542114*, 144. http://doi.org/10.1145/2046396.2046416

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2015, June 19). Power to the People: The Role of Humans in Interactive Machine Learning. American Association for Artificial Intelligence.

Basu, C., Yang, Q., Hungerman, D., Singhal, M., & Dragan, A. D. (2017). Do You Want Your Autonomous Car to Drive Like You? In *HRI '17: The Twelfth ACM/IEEE International Conference on Human Robot Interaction*. Vienna, Austria.

Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: Current issues and guidelines. *International Journal of Medical Informatics*, *77*, 81–97. http://doi.org/10.1016/j.ijmedinf.2006.11.006

Brownlee, J. (2015). Apple Finally Learns AI Is The New UI. Retrieved December 25, 2015, from http://www.fastcodesign.com/3047199/apple-finally-learns-ai-is-the-new-ui

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In *CHI '17 Proceedings of the 2017 annual conference on Human factors in computing systems*.

Fallman, D. (2003). Design-Oriented Human-Computer Interaction. In *Proceedings of the International Conference on*

Human Factors in Computing Systems (CHI'03) (pp. 225–232). http://doi.org/10.1145/642611.642652

Fiebrink, R., Cook, P. R., & Trueman, D. (2011). Human model evaluation in interactive supervised learning. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11* (p. 147). New York, New York, USA, NY, USA: ACM Press. http://doi.org/10.1145/1978942.1978965

Kay, M., Patel, S. N., & Kientz, J. A. (2015). How Good is 85%?: A Survey Tool to Connect Classifier Evaluation to Acceptability of Accuracy. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2015)* (pp. 347–356). http://doi.org/bqd5

Kuang, C. (2013, August). Why a New Golden Age for UI Design Is Around the Corner. *Wired Magazine 21.09*. Retrieved from http://www.wired.com/2013/08/design-and-the-digital-world/

Leahu, L. (2016). Ontological Surprises: A Relational Perspective on Machine Learning. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16* (pp. 182–186). New York, New York, USA: ACM Press. http://doi.org/10.1145/2901790.2901840

Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web - WWW '14* (pp. 677–686). New York, New York, USA: ACM Press. http://doi.org/10.1145/2566486.2568012

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype.

Patel, K. (2010). Lowering the barrier to applying machine learning. In *CHI '10 Doctoral Consortium* (pp. 355–358). http://doi.org/10.1145/1753846.1753882

Rader, E., & Gray, R. (2015). Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (pp. 173–182). New York, New York, USA: ACM Press. http://doi.org/10.1145/2702123.2702174

Schon, D. A., & DeSanctis, V. (1986). The Reflective Practitioner: How Professionals Think in Action. *The Journal of Continuing Higher Education*, *34*(3), 29–30. http://doi.org/10.1080/07377366.1986.10401080

Wood, M. (2014). Beats Hopes to Serve Up Music in a Novel Way. Retrieved from http://nyti.ms/1fr8LY5

Yang, Q., Zimmerman, J., & Steinfeld, A. (2015). Review of Medical Decision Support Tools : Emerging Opportunity for Interaction Design. In *IASDR 2015 Interplay Proceedings*. Brisbane, Australia. http://doi.org/10.13140/RG.2.1.1441.3284

Yang, Q., Zimmerman, J., Steinfeld, A., Carey, L., & Antaki, J. F. (2016). Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (pp. 4477–4488). New York, New York, USA: ACM Press. http://doi.org/10.1145/2858036.2858373

Yang, Q., Zimmerman, J., Steinfeld, A., & Tomasic, A. (2016). Planning Adaptive Mobile Experiences When Wireframing. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16* (pp. 565–576). New York, New York, USA: ACM Press. http://doi.org/10.1145/2901790.2901858

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research Through Design as a Method for Interaction Design Research in HCI. In *CHI '07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 493–502).

Zimmerman, J., Tomasic, A., Simmons, I., Hargraves, I., Mohnkern, K., Cornwell, J., & McGuire, R. M. (2007). Vio: a mixed-initiative approach to learning and automating procedural update tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07* (p. 1445). New York, New York, USA: ACM Press. http://doi.org/10.1145/1240624.1240843