ML for UX? - An Inventory and Predictions on the Use of Machine Learning Techniques for UX Research

Michael Chromik
LMU Munich
Munich, Germany
michael.chromik@ifi.lmu.de

Florian Lachner*
LMU Munich
Munich, Germany
florian.lachner@ifi.lmu.de

Andreas Butz
LMU Munich
Munich, Germany
butz@ifi.lmu.de

ABSTRACT

Machine learning (ML) techniques have successfully been applied to many complex domains. Yet, applying it to UX research (UXR) received little academic attention so far. To better understand how UX practitioners envision the synergies between empathy-focused UX work and data-driven ML techniques, we surveyed 49 practitioners experienced in UX, ML, or both and conducted 13 semi-structured interviews with UX experts. We derived an inventory of ML's impact on current UXR activities and practitioners' predictions about its potentials. We learned that ML methods may help to automate mundane tasks, complement decisions with data-driven insights, and enrich UXR with insights from users' emotional worlds. Challenges may arise from a potential obligation to utilize data and a more restrictive access to user data. We embed our insights into recent academic work on ML for UXR and discuss automated UX evaluation as a promising use case for future research.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; *User studies*; *Usability testing*.

KEYWORDS

User experience research; UX research; machine learning

ACM Reference Format:
Michael Chromik, Florian Lachner, and Andreas Butz. 2020. ML for UX? - An Inventory and Predictions on the Use of Machine Learning Techniques for UX Research. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20), October 25–29, 2020, Tallinn, Estonia.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3419249.3420163

1 INTRODUCTION

In recent years, many enterprises shifted their priorities from purely focusing on efficient production and distribution to creating memorable customer experiences. In this way, they hope to differentiate themselves from competitors and establish a competitive edge. This shift towards an *"experience economy"* [33] made the *user experience* (UX) a primary design goal. The term UX refers to *"a person's perceptions and responses that result from the use or anticipated use of a product, system, or service"* [4]. UX provides a holistic perspective and encompasses a person's emotions, feelings, and thoughts that may be formed before, during, or after the interaction [25, 35].

Building on this notion, the discipline of *UX research (UXR)* aims to understand and design people's experiences from end to end. UXR has emerged as an interdisciplinary field with influences from various disciplines such as cognitive science, psychology, and engineering. Each discipline contributes different terminologies, methods, and technologies to it. UX researchers frequently utilize qualitative methods, such as semi-structured interviews or surveys, while data-driven quantitative approaches are currently still less common [32]. The rare use of data-driven approaches by UX researchers is surprising, given the increasing data volumes in many contexts. Especially bigger enterprises increasingly compete in a data-driven environment and try to embrace the *"age of analytics"* [15].

Fueled by the availability of large data sets and affordable computing resources, *machine learning* (ML) methods have successfully been applied to complex problems in various domains. Historically, most academic research on ML within the HCI community had a technical focus on how to improve the interaction with systems (e.g., through adaptive interfaces) or develop new modes of interaction (e.g., gesture and voice interfaces) [6]. In the opposite sense, HCI academics have started to investigate how designers can enhance the user experience of ML-powered intelligent systems (*"human-centered machine learning"*) [11, 24] and how to address the distinct challenges of *human-AI interaction* [1, 46] from a *user-centric* perspective. However, there is little academic discourse that takes a *UX practitioner-centric* perspective and examines how ML methods could be leveraged to enhance the UX activities themselves. This lack of discourse may result from ML not yet being a standard part of the UX design practice as no relevant design patterns or prototyping tools have emerged yet [6]. Even if UX practitioners had previous exposure to ML, they often miss opportunities to use it in their design practice [45]. A review of HCI literature that employs ML observed that academics frequently resort to convenient interaction and design choices [44]. Thus, there may be a lack of awareness that the actual UX research and design processes may also benefit from ML.

To better understand the perception in the field, we have focused on practitioners to identify promising directions for the application of ML to UXR. We followed a two-pronged approach consisting of two independent studies to complement our insights from multiple angles. We surveyed 49 practitioners from the fields of ML and UX. Furthermore, we conducted 13 semi-structured interviews with

*now at Google

NordiCHI '20, October 25–29, 2020, Tallinn, Estonia

ACM ISBN 978-1-4503-7579-5/20/10…$15.00
https://doi.org/10.1145/3419249.3420163

UX practitioners who were educated or had experiences at the intersection between UX and ML. Our work contributes to the HCI research community in two ways: (1) We provide insights from two studies on ML's impact on current UX practices and ML's potentials for UXR. (2) We present data-driven UX evaluation using ML as a promising direction for future research and link it to recent academic work.

2 RELATED WORK

2.1 Terminology When using ML for UX

In the so-called *"experience economy"* [33] people use technology not only to accomplish a given task (i.e., for its *pragmatic quality*), but also to enjoy doing so (*hedonic quality*) [12]. The combination of both qualities constitutes the *user experience* (UX), i.e., the overall quality of a human's interaction with an interactive system. The field of UX covers an entire spectrum between the investigation to find user problems worth being addressed (*UX research*) and the creation of relevant interactions that provide a specific experience (*UX design*) [20].

ML refers to *"a set of methods that can automatically detect patterns in data [...] to predict future data, or to perform other kinds of decision making under uncertainty"* [34]. ML methods have successfully been applied to complex problems in a variety of domains such as spam detection, speech recognition, autonomous systems, and games. From a technical perspective, ML is typically split into *supervised* learning methods, which focus on predictions based on labeled training data, *unsupervised* learning methods, which find relationships in unlabeled data, and *reinforcement learning*, which optimizes some notion of reward by interacting with an environment. *Generative learning* methods create new contents such as texts or images. Approaching ML from a user-centered perspective, Yang et al. distinguish four channels of how it might generate value for users: inferring insights about an individual user, inferring insights about the context and external world (e.g., time, place, or social connections), inferring knowledge about how to optimize some arbitrary metric, and enabling entirely new user capabilities (utility) [44].

Combining the practices of UX and ML may yield positive effects in both directions: On one hand, knowledge in many domains is not only encapsulated in data, but also in the implicit expertise of human domain-experts. UX practice plays a key role in designing interfaces for those experts to effectively teach an ML model. In this way, UX decisions may have an impact on the model performance and robustness in the field (*interactive machine learning*) [7, 28]. On the other side, conversational UI and other forms of intelligent user interfaces offer new possibilities for UX design. Some observers claim that ML might become the most important design element to enhance user experiences by automatically personalizing systems to users and contexts (*"ML is the new UX"*) [47]. However, it has been observed that UX practitioners face challenges in understanding the data dependencies of ML and lack the tools to properly prototype with it [6, 43, 47].

2.2 Using ML for UX Research

Our work focuses primarily on *UX research* side of the spectrum. The goal of UXR is to systematically gather and analyze user data to understand a problem space and guide the entire design process. It is primarily applied at the generative and evaluative stages of the design process [9]. In the context of this paper, we subsume all empirical activities conducted by practitioners along the UX design spectrum as UXR. Building on the user-centered value channels of ML, using ML for UXR broadly falls under the *utility* channel [44]. ML indirectly benefits users through an improved UX if UXR practitioners can more effectively identify and validate user needs. A structured literature review by Yang et al. revealed that there is only little academic work at the intersection of UX and ML [44]. We found even less research that explicitly addresses ML for UXR. However, we noticed that the number of relevant publications has been increasing since 2015 and we expect that it will most likely continue to do so as ML is gaining popularity in many contexts. Below, we present some notable exceptions without claiming to be exhaustive.

Unlike conventional UXR approaches, that primarily generate new study data (e.g., through surveys or interviews), ML-based approaches were primarily used to enrich already collected user data. Most of this work analyzes *textual user data*. ML and natural language processing (NLP) methods have been used to semi-automate the coding of interview transcripts [29] and to extract UX-related problems from online review narratives through classification [13, 30, 40]. Data-driven learning approaches have also been used to construct *behavioral personas* derived from user clickstreams [48] and social media [19], and automatic real-time evaluation of usability and user experience via *emotional logging systems* using video-captured facial expressions in lab contexts [37] and on mobile devices [8], using acoustic data [36], and skin conductance signals [27]. Furthermore, ML was used for *selecting participants* for usability tests [10] and A/B tests [22].

2.3 Opposing Mindsets in UX and ML

Research on UX and ML originates from different academic communities. The relationship between the academic communities of HCI and AI has been discussed for decades. They tend to differ in their views of how humans and computers should interact with one another. These views can be roughly depicted along a spectrum of decreasing autonomy. While the HCI community values simplicity and user control, the sub symbolic fraction of the AI community favors the power of data-driven inference and convenience for the user. Winograd [42] argues that these views result from an opposing understanding of people and how technology is created for their benefit. He distinguishes two opposing approaches that exist in both communities. The *rationalistic approach* tries to depict the world through a quantitative or formal logic and tries to optimize the interaction accordingly. In contrast, the *design approach* acknowledges the complexities of the human world and tries to account for the limitations of modeling it. Instead, this approach focuses on the pragmatic interaction between a human and her or his environment.

Similarly, the UX mindset emphasizes the exploration of the desired future to be designed (design approach) while the ML mindset settles to accurately predict the future given data from the past (rationalistic approach) [43]. Opposing mindsets are also prevalent on the UX practitioners' side. UX has become increasingly cross-functional. Nowadays, many enterprises consider UX an organization-wide priority. This blurs the disciplinary boundaries between designers, developers, and marketers. UX teams often consist of experts from different disciplines [18]. UXR activities are seldom bundled in one role but often shared across the UX team. In practice, these teams must often cater to the needs of stakeholders with different mindsets: Colleagues with a design focus appreciate deep qualitative insights generated through user involvement. Additionally, business counterparts request aggregated quantitative insights to confirm their strategic decisions [26].

3 ONLINE SURVEY

3.1 Participants, Data Collection, and Analysis

With our survey, we intended to illuminate the impact of opposing mindsets on product development with a broader audience of ML and UX professionals. We were specifically interested in the differences between UX practitioners with and without experience in ML. In the last part of the survey, we examined if and how UX practitioners envision ML can be leveraged specifically for UXR activities. We designed a non-probabilistic self-selected survey that targets practitioners who have at least experience in either UX or ML, ideally both (inclusion criteria). Because the boundaries of UXR are fluid along the UX spectrum, we addressed a broader audience of UX professionals. We also assumed that ML developers are often involved at some stages of the UX process and could thus contribute valuable perspectives. The questionnaire consisted of 6 closed questions with ordered response options, 16 closed questions with unordered options and 4 open-ended questions. To understand the practitioners' contexts, we inquired about their demographics, educational background, working position and experience, and the qualitative and quantitative methods they apply regularly. Furthermore, we asked them to express their interpretation of the intersection between ML and UX and potential use cases for it. This way, we implicitly examined whether they could imagine potentials for UXR use cases. In the last part, we explicitly asked how they assessed the applicability, feasibility, and value of applying ML to different UXR use cases. The survey was designed according to the guidelines of the local institutional review board (IRB).

We pre-tested the survey with a few subjects to eliminate potential ambiguities and design flaws. We evaluated and incorporated their feedback into the final survey design. The survey was distributed through UX- and ML-related mailing lists of academic institutions in the United States and Germany as well as practitioner-oriented social media groups. Survey participants were self-selected and submitted their responses anonymously. As a reward for their participation, all respondents had the chance to take part in a lottery of three e-commerce vouchers worth 150 USD and two vouchers worth 60 USD. The survey was open for 4 weeks. 124 participants started the survey during this period. 19 participants did not meet the inclusion criteria. 56 participants did not finish the survey. After cleaning the data, we obtained 49 complete responses that met the inclusion criteria.

Respondents' demographics were quite diverse. 14 respondents self-identified themselves as female and 35 as male. Respondents are located in Germany (28), the United States (12) and other countries (9). 36 are working in the industry, 4 in academia, and 9 at the intersection of both. Their average work experience was 5.8 years (min=1, max=23 years). 17 respondents self-reported they have working experience only in UX (*UX-only*), 23 in UX and ML (*UX+ML*), and 9 respondents only in ML (*ML-only*). Most of UX-only respondents described their primary role as UX designer or UX consultant, UX+ML respondents as product manager, UX designer or UX researcher, and ML-only respondents as ML engineer/developer. 13 of the 17 UX-only respondents assessed their knowledge of ML as *basic* (familiar with the term and basic concepts) while 16 of the 23 UX+ML respondents consider their knowledge of ML as *advanced* (basic practical experiences) or *expert-level* (applied experience in the field of ML). All ML-only respondents assessed their knowledge as advanced or expert-level. In total, 3 respondents stated they are unfamiliar with ML (all in UX-only).

3.2 Findings

Our analysis of responses indicates that UX practitioners with ML experience have a different take on UX and more often leverage quantitative methods as part of their daily work. Most of the respondents believe that ML and UX will increasingly overlap in the future. Lastly, respondents consider the data-driven evaluation of UX as a promising use case for applying ML to UX research.

3.2.1 Current Project Involvement and Research Methods. Most of the respondents are involved in the pre-deployment stages of product development. There, the respondents work mainly on the initial development (e.g., wireframing, low-fidelity prototyping) and final development (e.g., high-fidelity prototyping, final product) of a product. In our sample, we see a trend that UX-only respondents are more often responsible for the conceptual stages such as product vision development or needs research (88% for UX-only compared to 43% for UX+ML). In contrast, UX respondents with ML experience are slightly more often involved in the final development stages (87% compared to 71%). Only half of the respondents (27 out of 49) are regularly involved in the evaluation of a product after its launch.

Overall, 30 out of 49 (61%) respondents apply qualitative and quantitative methods equally often as part of their daily work. However, this is only the case for 8 out of 17 (47%) UX-only respondents. 7 of them are mainly qualitative researchers. In contrast, 16 out of 23 UX+ML respondents (70%) apply both types of methods equally often. This trend is also reflected in different opinions on how UX should be approached. We asked participants about their agreement using a 6-point Likert-scale (1=disagree very strongly, 6=agree very strongly). 76% of UX-only respondents agree or agree strongly that UX must be approached qualitatively (compared to 57% of UX+ML respondents). Furthermore, 65% of UX-only respondents agree or agree strongly that UX can be quantified (compared to 87% of UX+ML respondents).

Respondents mostly agree on when to use qualitative methods. For qualitative methods, we observe large differences between UX-only and UX+ML respondents. When talking to their project stakeholders, half of the respondents argue with qualitative insights. However, 61% of UX+ML respondents leverage quantitative data to back their arguments while only 29% of UX-only respondents

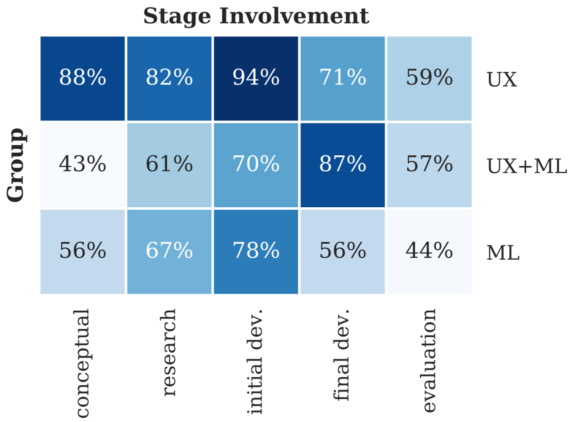

Figure 1: Respondents' involvements in the product development stages per group based on 49 respondents (17 UX-only, 23 UX+ML, and 9 ML-only). Each cell represents the percentage of respondents in the group that stated to be involved in the respective stage.

do so. Similarly, we observe differences when choosing between design options. Proportionally twice as many UX+ML respondents leverage quantitative methods to back their decisions in addition to qualitative methods (18% UX-only vs. 43% UX+ML). When it comes to individual research methods, such as semi-structured interviews or questionnaires, we see roughly equally often usage. Yet, we see a difference in leveraging user log data and online feedback. 35% and 41% of UX-only respondents apply these methods in at least half of their projects, respectively (compared to 74% and 74% of UX+ML respondents, respectively).

3.2.2 ML and UX Are Expected to Overlap in the Future. Respondents were asked to assess their current perception of the interplay between ML and UX, and how they predict it will evolve in the future. Assessing the status quo, only 9 respondents perceive ML and UX to overlap to some or great extent. However, 23 expect ML and UX to overlap at least to some extent in the future. In total, 35 out of 49 respondents think that the interplay between both disciplines will increase in the future. None of the respondents are expecting that the disciplines will drift apart (see Figure 2).

Next, we asked respondents to describe the perceived interplay in their words. We asked them to illustrate it based on a promising scenario from their daily work. We aimed to examine their perception of applying ML to UXR without directly asking them about it. We grouped the mentioned scenarios by use case: 19 respondents mentioned use cases that aim to improve the UX of products for users through ML features, 17 mentioned use cases that enhance UX research and design activities, 11 mentioned use cases about improving the UX of developing ML models, and 8 mentioned miscellaneous use cases (some respondents described multiple use cases). The UX research and design use cases included the use of ML *to reveal user insights* (6 mentions; e.g., trends in user behavior, analysis of user reactions, identifying plots in user study results), to *evaluate the UX of products* (6 mentions; e.g., automated measurement of UX, predicting the UX of new users, continuous UX monitoring) and *to augment the creation of UX artifacts* (5 mentions; e.g., producing variations of interaction flows, automating standard design tasks, informing design with historical data).

3.2.3 Leveraging ML for UX Research. Since we were interested in the opportunities for applying ML to UXR, we subsequently asked all participants about activities and use cases that are specifically related to UXR. We asked them to assess the potential per use case taking applicability, feasibility, and value into account. Respondents consider applying ML to yield insights from log and time series data, to remotely track user behavior over time, and user modeling as promising fields for future exploration (see Figure 3).

Furthermore, we asked which types of ML they had in mind when assessing the use cases: (1) prediction of an outcome based on the analysis of given data, (2) detection of patterns within structured or unstructured data, (3) generation of new outcomes or data, or (4) other. Most respondents think of scenarios for pattern detection and outcome prediction. UX-only respondents are more optimistic about the potentials of generative learning approaches. 41% of UX-only respondents consider them feasible and valuable. In contrast, UX+ML (13%) and ML-only (11%) respondents are much more conservative.

Next, we questioned for which stages of the product development process they perceive ML to be well-suited. We provided them with typical example activities for each stage. Respondents think that ML is especially applicable to later stages of the development process. 41 out of 49 believe ML can be applied to some or to a great extent to evaluate and test products after their development (e.g., UX evaluation of products on the market). On the other hand, few respondents can envision how ML can support the conceptual stage of product development (e.g., product vision or strategy development). The opinions tend to be divided for the stages of research (e.g., user research) and initial development (e.g., wireframing or prototyping) (see Figure 2). When comparing the results between the three groups, we observe that UX-only respondents have almost equal assessments for the first four stages. In contrast, UX+ML respondents have a more distinguished opinion. They see more potential in the initial as well as final development stages. The assessment of the UX+ML respondents is very much in unison to the assessment of the ML-only respondents.

4 EXPERT INTERVIEWS

4.1 Participants, Data Collection, and Analysis

We conducted semi-structured interviews with 13 UX experts from industry and academia to understand how they envision ML methods to enhance or influence their UX processes. We recruited experts from the fields of Human-Computer Interaction (HCI) and UX who are experienced with the concepts and applications of ML (either by professional collaboration with ML engineers or by education). As the intersection of UX and ML is a young field, we aimed for a mix of experienced senior professionals as well as young professionals (who were trained in both fields as part of their study program). Starting the recruiting through our academic network, we asked each participant to recommend experts who potentially meet our criteria for further interviews (*snowball sampling*). Our panel comprised mostly UX professionals from leading digital companies as

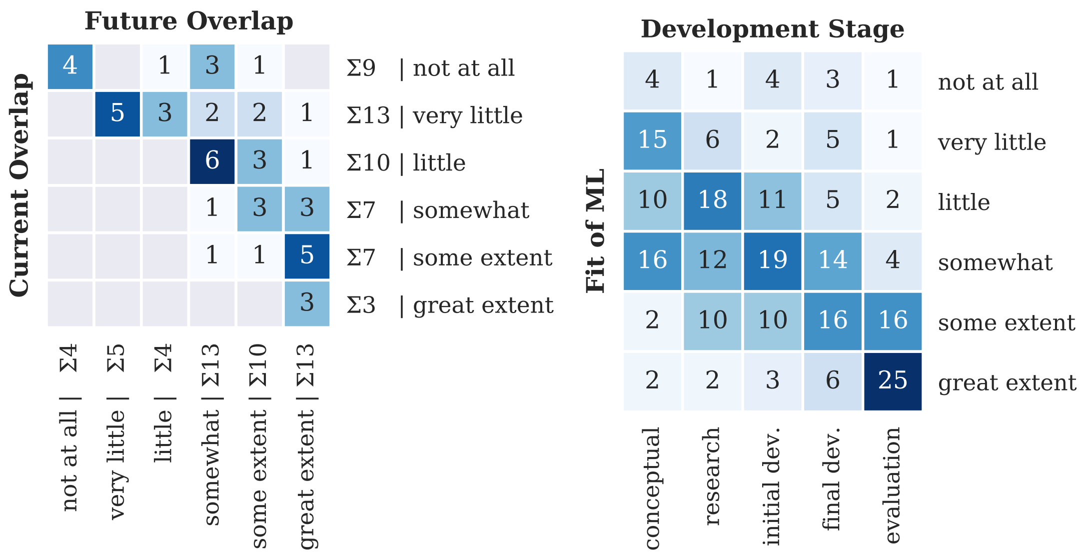

Figure 2: *(Left)* Perceived overlap of ML and UX aggregated by the number of responses (N=49). Entries on the diagonal mean that no change is expected. Entries towards the upper right corner mean that the overlap is expected to increase (e.g., of the 13 respondents that currently see *very little* overlap, 8 respondents expect the overlap to increase in future). *(Right)* Respondents' assessment of how well ML techniques can be applied to the respective stages of the product development process (N=49 for each development stage).

well as graduates from a relevant interdisciplinary study program at a renowned academic institution. Table 1 presents a summary of the participants' characteristics. The interviews were conducted in person or via video calls and lasted roughly forty-five minutes each. The sessions were recorded and transcribed to analyze them further (total recording time of 12 hours).

To understand the contexts of the participants, we asked them about their backgrounds, work routines as well as the importance of UX within their institution. We inquired about previous projects in which they had applied data logging to get a sense of their exposure to quantitative research methods and ML. Furthermore, we asked them to ideate how ML might enhance their UX method toolbox or enable novel ways of UX research. We asked them to ideate around a hypothetical ML system that automatically evaluates the UX of a user during interaction with a product based on usage data. Lastly, we asked them what challenges they thought stakeholders in the UX research process might face when applying ML-based methods, especially in terms of privacy and ethics.

For data analysis purposes, we transcribed the audio recordings from the expert interviews. Then we followed a *Grounded Theory*-inspired emergent coding approach, i.e., we analyzed without a guiding theory in mind. In a first step, one author extracted 120 UX and ML-related trains of thought from the interviews (each consisting of one to many sentences) and gave each observed phenomenon a distinctive name using mostly in-vivo codes. The author also identified connections between the codes and grouped them in multiple iterations into higher-level themes. Those themes are represented by the derived opportunities and challenges. In a second step, two authors independently coded the extracted trains of thoughts given the codebook of higher-level themes from step one. The inter-rater reliability was $\alpha = .8710$ with 95% confidence in a CI of (0.8015, 0.9305). According to Krippendorff [23], values for α greater than .8 can be considered satisfactory. Typically, Grounded Theory (GT) starts from a set of empirical observations and aims to reverse-engineer a hypothesis from the observations in multiple iterations [38]. Our approach follows the GT approach in terms of open and axial coding. However, we are not formulating a (well-grounded) theory from our observations as we primarily aim to describe and group practitioners' opinions in terms of perceived challenges and opportunities.

4.2 Findings: Opportunities

Our analysis revealed 3 areas of opportunity along the dimensions of *automating*, *complementing*, and *enriching* the insight generation practices of UX researchers.

4.2.1 Automate the Mundane Part. ML is often perceived as a tool to free people from time-consuming and repetitive tasks of limited value. In this sense, our participants saw opportunities for ML to (semi)-automate parts of their current work routines. Furthermore, designing survey studies as a more engaging and personalized experience could result in richer user data and higher response rates.

Automated Transcription: ML-based speech-to-text services were hoped to significantly shorten the time between data collection and data analysis. This was considered particularly interesting for

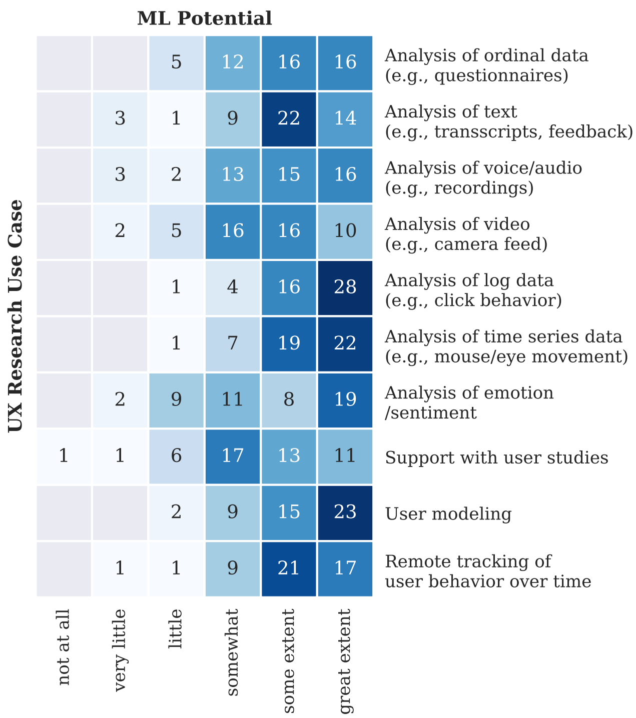

Figure 3: The distribution of respondents' assessments regarding the potentials for applying ML to UXR use cases taking applicability, feasibility, and value into account (N=49 for each use case).

the post-processing of contextual inquiries or interviews since *"you don't need to transcribe or code all the information"* (P11). Instead, *"you can give them [the ML systems] your audio file and they'll send you a typed-up version of it."* (P6). *"Instead of trying to scribble notes or record the conversation and then transcribe it, it's just doing it while you're in the field."* (P10). When ML takes over mundane tasks, UX researchers are enabled to focus on higher-value activities. *"AI can do things to make us faster at producing the kind of work that we want to do versus the kind of work we have to do"* (P1).

Engaging Surveys: Participants felt that ML can simplify survey studies for researchers and survey participants alike by leveraging the idea of adaptive user interfaces. Questionnaires might automatically be tailored in real-time to the individual survey participant based on their previous answers as well as answers of similar participants. The intend is *"to not give everybody 100 questions, but just give the 25 most important ones"* (P5). Also, advances in conversational and voice user interfaces were considered an opportunity for more empathetic survey studies. A questionnaire might be turned into engaging conversational experiences by *"acting like it's a person, giving it a personality"* (P10). Thereby researchers would receive *"a response based on a conversation rather than just filling out a survey question"* (P10). One participant described an industry project where the questionnaire mimicked a TV personality to better engage with teenagers. Voice user interfaces could furthermore enrich the responses with affective signals derived from speech.

4.2.2 *Complement With Undrawn Data.*

ML methods excel at quickly analyzing vast amounts of existing data. Leveraging this capability, participants see opportunities for ML to identify subtle patterns in dispersed data silos as well as to inform UX decisions with data insights.

Linking Insights Across Data Silos: Participants believed that ML can augment UX researchers to *"understand the links between data sources"* (P13) and see *"if there are any behavioral patterns, [or] pain points that we overlooked during the quantitative analysis."* (P5). Participants envisioned that with the help of ML tools they could *"map people to other data sets that we have"* (P13), such as clickstreams, social media, similar interviews, or survey responses. Doing such analyses manually is often time-consuming and slows down the line of thought, thus their potentials remain currently untapped. Participants perceived that ML methods might broaden their scope while leaving the interpretation with the human. *"Whenever you look at information just from one data set - it's like shining the flashlight only in one corner. [...] ML can help us to illuminate more."* (P7). *"It's going to be helpful to understand the bigger picture. [...] It's going to be quicker. [...] At the moment I don't see much use of AI to help us to understand the why"* (P11).

Data-Driven Personas: Furthermore, many participants saw potential for unsupervised ML in supporting user segmentation. Clustering methods may automatically identify unique user groups from data logs. *"Don't make me do all the work to figure out what kind of user segments I have"* (P1). Instead, tools might provide analytical insights on how many clusters can be observed in the data and let the UX researcher fill in the details. This could also help to *"keep the user researcher unbiased"* (P1). Unsupervised ML methods often identify patterns for which *"there is not really a human word"* (P4) and challenge researchers' potentially biased assumptions. Additionally, a data-driven persona approach could enable UX researchers to monitor user shifts more closely over the product life cycle. Currently, personas are often created once in the beginning *"and maybe you do it again in a couple of years"* (P4). UX researcher might be notified when significant changes are observable in transaction data that require adaptation of personas.

Data-Driven Design: Supporting design decisions by evaluating and recommending UI options based on historical data of user behavior or user preferences was considered another field of interest. *"We can use ML and its potential to help make good decisions in design"* (P7). P3 would like to see data-driven design tools that back the UX design process with actual numbers, e.g., *"with this particular design this is the problem [...] because 50% of the users failed at this particular step."*

4.2.3 *Novel Insights into Users' Affect.*

In addition to improving current practices and connecting existing data, participants envisioned ML to yield novel information about users' feelings and emotions. This would enable UX researchers to better *"understand the affective component"* (P9). *"The one thing I still don't have access to is sentiment. I don't know their emotional state. Often a system can figure out the emotional state [of a user] faster [than humans]"* (P1).

Generalizing Beyond the Lab: Intrusive methods, such as electroencephalography (EEG) sensors, could be used during real-time usability tests in lab contexts to record typical flows of interaction and their corresponding emotional responses. P4 states that their behavioral and emotional responses could be used as labels for an ML model that could then be applied to *"other users in mass"*.

Non-intrusive and Remote: Affective signals could also be derived from remote test settings to provide UX researchers with richer context when studying product use in the field. Specifically trained ML models could provide *"a better window into the emotional state that people are actually feeling"* (P9) through *"non-intrusive measures of sentiment"* (P1). For example, P1 envisioned ML to continuously classify users' facial expressions from an *"accompanying web camera feed"* and reveal that a user *"was actually looking over here and was chatting with his wife"*.

Identifying User Inconsistencies: People are sometimes observed to provide inconsistent feedback, i.e., users may *"say one thing but do another"* (P9). Affective signals may be compared to the actual behavior and thus help UX researchers to identify inconsistencies. *"These mechanisms could help uncover some of the usability flaws that are very difficult to extract with manual methods"* (P2).

Virtual User Testing: Further down the road, P9 saw potential in conducting UX studies entirely in a virtual setting with the help of virtual reality (VR). A virtual world could resemble the physical world but enable researchers to stimulate responses that are hard to simulate in the physical world. In such a virtual environment, ML methods could be used to evaluate eye gaze, motion, and neurological activity when people are experiencing those situations.

4.3 Findings: Challenges

Furthermore, we identified 2 emergent areas of challenges. First, participants foresee changing expectations towards the UX profession and a shift in future responsibilities. Second, ML was seen to make it more difficult to recruit human subjects for UXR activities.

4.3.1 Changing Expectations and Responsibilities. Participants felt that the availability of data might oblige them to report quantified insights while not feeling entirely prepared for it. Furthermore, some participants perceived that ML changes how UX researchers will be involved in projects.

Peers Demand Numbers: Driven by the promises of ML, our participants felt that leveraging large-scale data for UXR might be increasingly demanded by their peers. The potential availability of data might make expressing insights through aggregated numbers mandatory. P5 described cases where it was necessary to use quantified insights *"to convince product managers or management because without numbers it's oftentimes very hard to get somebody to understand what is happening. We already have this but need numbers."* P7 explained that *"having numbers makes it feel more scientific, even though that's not necessarily the case. [...] It's kind of a pervasive problem in [the] industry that people think only numbers are true."* A key challenge in our participants' view was *"how to balance [those] different analytical needs"* (P3). While most people in an organization require an aggregated view to understand the bigger picture, UX researchers cherish to *"look at individual flows"* (P3) to address underlying problems. *"Any good [UX] researcher or good [UX] designer would start with a user need"* (P1). Data alone leaves many interpretations. So, it is mandatory *"to enrich it with qualitative insights"* (P1). Convincing internal peers of the need for resource-intensive low-number qualitative insights might become more challenging as ML is successfully applied to other parts of an organization. *"Qualitative data is only as powerful to those who participate in it and can see the actual results. [...] Most people aren't trained to understand this thing I call qualitative validity"* (P7). To advocate the validity of qualitative insights might become more challenging for UX practitioners when not supported by numbers.

Developing Confidence and Literacy in ML: Figuring out what to do in a world of data streams was perceived as a complex challenge for UX teams. *"The problem with data is that there is so much of it. The world [...] becomes even more complex because all those data streams don't go away"* (P1). Participants perceived that UX researchers *"are not completely educated about ML [...] and do not understand that the two can work together"* (P5). Participants admitted that a cultural change is needed among UX practitioners to foster a data-driven spirit in organizations. *"A lot of user researchers are essentially traditional qualitative researchers. There is a little bit of resistance [...], but that's becoming lesser and lesser given that management wants it to be both [qualitative and quantitative]"* (P5). On the technological side, participants observed that ML remains an inaccessible design material as the usability of ML tools is often a hurdle for UX practitioners. UX teams must blindly trust the default settings of tools as they do not understand the technical foundations. *"ML is totally inaccessible to anyone who has never coded. People just trust these out of the box models and try to get it to work. It's not [going to]."* (P4). P4 would appreciate less technical terms in ML tools. Instead, the participant would like to *"call it what it is, like tell what problem it is solving"*. In-browser ML environments and interactive ML approaches have been named as examples. Often there seems to be a common belief within organizations that people could simply run ML on a problem and would obtain a meaningful solution. *"A lot of problems aren't scoped in a way that ML can help"* (P4). Interpreting the potentials of ML and having the vocabulary and confidence to argue about it with stakeholders was perceived as an obstacle for current UX practitioners.

Changing Project Involvement: Overall, participants expressed little concern that applying ML methods to UXR would reduce the demand for human researchers. *"Qualitative methods [...] result in rich data, that is only truly understandable by [...] a human being. There is so much information [...] that a machine would have a very hard time truly understanding it. It requires actual empathy and cultural appreciations"* (P7). *"I don't think any machine will ever get to the point where we trust the AI more than we trust the [UX] person"* (P1). However, there was some disagreement among participants about how the skill set of UX researchers might change in the future due to ML. On one hand, some participants believed that the role of UX will likely stay the same. *"I don't think the skill set would change. You still need to do all the things [...] to understand human behavior"* (P11). In contrast, P4 believed that ML and data-driven methods are not only changing the mindset of UX researchers but *"how people are currently doing their jobs" (P4).* Working on ML-enabled products was considered an ongoing process that will involve researchers over longer periods before becoming effective for users. This contrasts with currently established design thinking approaches, where UX researchers tend to move on to the next project after few prototype iterations (P4). Some participants believed that UXR will become even more interdisciplinary. Other disciplines, such as anthropology and sociology, might increasingly

Table 1: Participants in the expert interviews including their role, country, institution size, and work experience in the field of UX. Participants P11 and P13 asked us to omit their work-related information.

Participant	Current Role	Country of Residence	Size of Institution	Work Experience in UX
P1	UX Researcher	USA	<50	20+ years
P2	Student	India	*did not disclose*	4 years
P3	Data Analyst/Scientist	USA	1,000+	1 year
P4	UX Designer	USA	1,000+	4 years
P5	UX Designer	USA	100+	1 year
P6	UX Designer	USA	100+	1 year
P7	UX Executive	USA	<50	25+ year
P8	UX Researcher	Germany	250+	3 year
P9	Academic Researcher	USA	1,000+	13 years
P10	UX Researcher	USA	50+	1 year
P11	*did not disclose*	*did not disclose*	*did not disclose*	*did not disclose*
P12	Academic Researcher	USA	*did not disclose*	4 years
P13	*did not disclose*	*did not disclose*	*did not disclose*	*did not disclose*

contribute to the study of complex human phenomena in collaboration with current disciplines. *"It's not one single skill set anymore that you apply to understand the users"* (P1). Instead, participant P1 envisioned UX researchers to *"become a translator"* between the stakeholders involved.

4.3.2 *Access to Users and Their Data.*

To effectively leverage ML methods, access to large amounts of user data is necessary. However, ML's reliance on data resulted in an increasing number of regulations and increased sensitivity regarding user data usage. Participants saw challenges in interpreting these regulations, balancing data economy, and finding alternative means to incentify users to participate in UXR activities.

Interpreting Privacy Regulations: Getting access to users was considered a major constraint for UX researchers as it imposes legal, confidential, and financial requirements. *"Recruiting [users] will always be the golden key"* (P1). New privacy regulations, such as the European Union *General Data Protection Regulation (GDPR)* [39], aim to improve the control for users over their data. Interpreting those regulations and finding the right balance between advocating in favor of users versus pursuing organizational interests was considered a major challenge for the time ahead. Some participants believed that UX researchers *"need to err on the side of [data] protection"* (P1) while others felt that *"any data can be used for analysis as long as PII [personal identifiable information] data is not used"* (P5).

Dealing with the Principle of Data Economy: We observed different opinions on the importance of individual data in user behavior tracking. Some participants think that the principle of data economy may limit their access to user data. Others feel that having access to aggregated data might be enough for most use cases. To understand the big picture, P8 perceived it to be more important *"to see the behavior of one average user than to watch individual cases"*. *"I want to know when it fails. That does not need to be tied to [...] username"* (P4). Instead of tracking everyone by default, UX researchers could also turn to selectively ask individual users for feedback, e.g., via pop-up surveys on a website (P5). Excluding demographic data from individual cases may even have positive effects in terms of bias avoidance. *"I had to keep telling myself that I can't be biased over some person's background since that kind of information is not available when we generalize"* (P5).

Incentivizing Users to Contribute: As an alternative way forward, some participants felt that UX researchers and companies should rethink their relationship with user data. *"We need to give a lot more credit to the producers of the data"* (P1). They envisioned ways to encourage users to contribute their data to UXR. P1 suggested some form of *"privacy currency"* that offers benefits, such as *"reduced number of ads"* or *"5% off the purchase price"*. Companies should be more honest about their need for usage data. *"Don't automatically opt everybody in. Give them the option. Make it easy. People appreciate that more than having to dig through layers and layers of UI to uncheck a box"* (P6). None of the participants reported practical experiences in this direction. While participants seemed positive about such alternatives to compensate for potentially fewer user data, these approaches also entail challenges in promoting and implementing them internally and externally.

5 DISCUSSION

The notion of the *fourth wave of HCI* [2, 3] speculates that HCI is converging towards a trans-disciplinary paradigm as new disciplines enter the stage. Each discipline adds new dimensions, such as ethics or creativity, to the interdisciplinary discourse. Our findings from the survey and the interviews suggest that the discipline of ML entered the UXR discipline even though it may not be effectively applied to UX practices yet. Furthermore, it shifts the mindsets and work practices of practitioners towards a more quantitative interpretation of UX. Our identified opportunities overlap with findings from prior research. [45] report how UX practitioners enrich their UXR toolkit through telemetry and data stories. Further, some of our themes resemble the user-centric perspective of ML by [44]. Their perspective describes how ML can provide direct value to the user by enabling them to understand themselves (e.g., through insights into their affect) or their surrounding (e.g., through insights across data silos). This assumes that this is done dynamically by the system without a UX researcher in the loop. Our findings suggest that ML may also be used to indirectly provide value to the user by informing UXR activities. Our identified challenges indicate that

Table 2: Summary of themes observed in the expert interviews.

Higher-level Theme	*Emergent Theme*
Opportunities	
Automate the Mundane Part	Automated Transcription Engaging Surveys
Complement With Undrawn Data	Linking Insights Across Data Silos Data-Driven Personas Data-Driven Design
Novel Insights Into Users' Affect	Generalizing Beyond the Lab Non-intrusive and Remote Identifying User Inconsistencies Virtual Testing
Challenges	
Changing Expectations and Roles	Peers Demand Numbers Confidence and Literacy in ML Changing Project Involvement
Access to Users and Their Data	Interpreting Privacy Regulations Dealing with Data Economy Incentivizing Users to Contribute

UX researchers' core skills of interpersonal communication are expected to advance beyond the focus on users. Instead, they translate between multiple stakeholders as well as privacy requirements. Weaving in insights derived from data-trails and ML techniques may be required to persuade stakeholders that their conclusions are valid and will solve a relevant problem. Based on the interpretation of our findings, we see three promising directions for further HCI research that have not yet been adequately addressed.

5.1 Data-driven UX Evaluation With ML

Our findings indicate that using ML for the evaluation of a product's UX may be a promising field for future research. Most of the respondents believe that ML can provide the biggest value at the evaluation stage. Traditional UX evaluation methods are often resource-intensive and not scalable. Often standardized questionnaires such as the *user experience questionnaire (UEQ)* or the *AttrakDiff questionnaire (AD)* are used [32]. ML techniques may offer a more resource-effective alternative. Connecting questionnaire results with log and time series data about user behavior may be used as labeled data for supervised ML. Furthermore, such approaches may allow to continuously monitor changes in users' UX and inform UX researchers when it might be worth to revisit parts of the product experience. We observed that fewer UX practitioners are involved in the evaluation of a product's UX after its launch. Thus, respondents' assessment' may be positively biased because they may not have a complete picture of potential obstacles in this field. However, we found recent academic work that explores the challenges of evaluating UX using multiple data sources and proposes ML-based approaches [17, 31]. We propose to explore sensitizing concepts for ML-supported continuous UX evaluation and UX monitoring in future work.

5.2 Ensuring Effectiveness of ML-based UXR

Insights from data-driven ML techniques have the potential for effective triangulation to ultimately yield a more complete picture [32]. This matches the opportunities identified by our study participants. However, they should be carefully evaluated in practice. Critical voices have been raised about the practical applicability of automated systems in the field. Previous works compared automated ML approaches for UX research with traditional (manual) methods [14]. Results indicated that issues extracted by algorithms might differ after deployment to the field – even though they looked precise during training. UX researchers need to be able to spot questionable predictions and develop an understanding of when to rely on automated methods and when to carefully supervise them. Building ML tools for UX activities around the guidelines for *interactive ML (IML)* [7] and *explainable artificial intelligence (XAI)* [41] may be promising directions to enable UX researchers to validate and maintain the effectiveness of such tools in the field.

5.3 Calibrating Expectations Regarding ML

UX practitioners have been confronted with many novel forms of technology and interaction. Multi-device experiences, voice interfaces, and unpredictable intelligent systems pose new challenges and opportunities in terms of UX research and design [6, 43, 47]. The HCI community already raised the question of whether current methods are keeping up with the technological advancements and user expectations [16]. In line with prior work, almost all our inquired UX practitioners experimented with the new design material ML at least on a basic level – even when their work practices may primarily be qualitative. However, we observed opposing mindsets between UX practitioners with and without ML work experience. The assessments of UX+ML respondents have often been in unison with ML-only respondents. We interpret this in a way that UX practitioners with work experience in ML have a sufficient understanding to realistically assess capabilities but also limitations of ML – even though they are no technical experts. In contrast, UX-only practitioners may envision more creative use cases, e.g., regarding generative approaches, because their knowledge about difficulties in practice is limited. This might imply that UX researchers would benefit from more distinguished educational material that also addresses ML's limitations. Recent academic work lets UX practitioners refine their mental models with tools for playful exploration [21] and metaphors [5]. We suggest that such educational materials also include case studies on how to apply ML to UXR use cases.

5.4 Limitations

We acknowledge that our findings are indications that can only be generalized to a limited extent. Our participants were not selected for demographically representative proportions. The studies recruited mainly participants from the United States and Germany and were limited in time. Further, we asked our participants to reflect on the potential of ML in the future. As 13 out of 49 participants (especially in the UX-only group) had only a basic understanding of ML, some future predictions might turn out to be too optimistic. Additional experts from adjacent disciplines should be interviewed and the derived insights should be related to our analysis. Still, we are confident that our studies capture up-to-date insights about practitioners' understanding and serve as an informative first step for future work in the emerging research field of ML for UXR. We welcome other researchers to extend or amend our insights and interpretations. Eventually, we will only be able to draw a complete picture of the applicability and acceptance of ML for UX when we conceptualize, develop, and evaluate respective tools and methods in case studies and prototypes.

6 CONCLUSION

The disciplines of ML and UX are contesting each other's borders. There is ongoing research within the HCI and UX communities on how to improve the performance of ML models through UX as well as research on how to use ML to improve a product's UX. With our work, we add the intersection of *ML for UX research* to the discussion. We found promising academic work that already experimented at the intersection of ML for UXR. Based on these, we surveyed and interviewed UX and ML practitioners. We presented practitioners' experiences and visions derived from a snapshot of 49 survey responses from UX and ML practitioners as well as 13 interviews with UX experts. Our survey indicated that the disciplines of ML and UX are expected to overlap and that UX practitioners see promising use cases of applying ML to UXR. Further, they are anticipating these developments as they are experimenting with ML even though their work routines may primarily be qualitative. We learned from the interviews that ML methods may help to automate mundane tasks, complement decisions with data-driven insights, and enrich UXR with insights from users' emotional worlds. We link our interpretations to recent academic work on ML for UXR and discuss data-driven UX evaluation based on ML as a promising use case for future research.

ACKNOWLEDGMENTS

We thank all interview partners and survey participants who took their time to contribute their experiences and opinions. The authors gratefully acknowledge the grant provided by the German Federal Ministry for Education and Research (#01IS12057).

REFERENCES

[1] Saleema Amershi, Daniel S. Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi T. Iqbal, P. Bennett, Kori Inkpen Quinn, J. Teevan, Ruth Kikin-Gil, and E. Horvitz. 2019. Guidelines for Human-AI Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019).

[2] Eli Blevis, Kenny Chow, Ilpo Koskinen, Sharon Poggenpohl, and Christine Tsin. 2014. Billions of Interaction Designers. *interactions* 21, 6 (Oct. 2014), 34–41. https://doi.org/10.1145/2674931

[3] Susanne Bødker. 2015. Third-wave HCI, 10 Years Later—participation and Sharing. *interactions* 22, 5 (Aug. 2015), 24–31. https://doi.org/10.1145/2804405

[4] ISO DIS. 2009. 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *International Standardization Organization (ISO). Switzerland* (2009).

[5] Graham Dove and Anne-Laure Fayard. 2020. Monsters, Metaphors, and Machine Learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376275

[6] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning As a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 278–288. https://doi.org/10.1145/3025453.3025739

[7] John J. Dudley and Per Ola Kristensson. 2018. A Review of User Interface Design for Interactive Machine Learning. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 8 (June 2018), 37 pages. https://doi.org/10.1145/3185517

[8] Jackson Feijó Filho, Thiago Valle, and Wilson Prata. 2015. Automated Usability Tests for Mobile Devices Through Live Emotions Logging. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Copenhagen, Denmark) *(MobileHCI '15)*. ACM, New York, NY, USA, 636–643. https://doi.org/10.1145/2786567.2792902

[9] Interaction Design Foundation. 2020. *What is UX Research?* Retrieved April 27th, 2020 from https://www.interaction-design.org/literature/topics/ux-research

[10] Juan E Gilbert, Andrea Williams, and Cheryl D Seals. 2007. Clustering for usability participant selection. *Journal of Usability Studies* 3, 1 (2007), 40–52.

[11] Marco Gillies, Rebecca Fiebrink, Atau Tanaka, Jérémie Garcia, Frédéric Bevilacqua, Alexis Heloir, Fabrizio Nunnari, Wendy Mackay, Saleema Amershi, Bongshin Lee, Nicolas d'Alessandro, Joëlle Tilmanne, Todd Kulesza, and Baptiste Caramiaux. 2016. Human-Centred Machine Learning. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. ACM, New York, NY, USA, 3558–3565. https://doi.org/10.1145/2851581.2856492

[12] Marc Hassenzahl. 2008. User Experience (UX): Towards an Experiential Perspective on Product Quality. In *Proceedings of the 20th Conference on L'Interaction Homme-Machine* (Metz, France) *(IHM '08)*. ACM, New York, NY, USA, 11–15. https://doi.org/10.1145/1512714.1512717

[13] Steffen Hedegaard and Jakob Grue Simonsen. 2013. Extracting Usability and User Experience Information from Online User Reviews. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 2089–2098. https://doi.org/10.1145/2470654.2481286

[14] Steffen Hedegaard and Jakob Grue Simonsen. 2014. Mining Until It Hurts: Automatic Extraction of Usability Issues from Online Reviews Compared to Traditional Usability Evaluation. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (Helsinki, Finland) *(NordiCHI '14)*. ACM, New York, NY, USA, 157–166. https://doi.org/10.1145/2639189.2639211

[15] Nicolaus Henke, Jacques Bughin, Michael Chui, James Manyika, Tamim Saleh, Bill Wiseman, and Guru Sethupathy. 2016. The age of analytics: Competing in a data-driven world. *McKinsey Global Institute* 4 (2016).

[16] Karen Holtzblatt, Ilpo Koskinen, Janaki Kumar, David Rondeau, and John Zimmerman. 2014. Design Methods for the Future That is Now: Have Disruptive Technologies Disrupted Our Design Methodologies?. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. ACM, New York, NY, USA, 1063–1068. https://doi.org/10.1145/2559206.2579401

[17] Md Sazzad Hossain, Amin Ahsan Ali, and M. Ashraful Amin. 2019. Eye-Gaze to Screen Location Mapping for UI Evaluation of Webpages. In *Proceedings of the 2019 3rd International Conference on Graphics and Signal Processing* (Hong Kong, Hong Kong) *(ICGSP '19)*. Association for Computing Machinery, New York, NY, USA, 100–104. https://doi.org/10.1145/3338472.3338483

[18] Christian Jetter and Jens Gerken. 2007. A simplified model of user experience for practical application. In *2nd COST294-MAUSE*. 106–111.

[19] Soon-Gyo Jung, Jisun An, Haewoon Kwak, Moeed Ahmad, Lene Nielsen, and Bernard J. Jansen. 2017. Persona Generation from Aggregated Social Media Data. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 1748–1755. https://doi.org/10.1145/3027063.3053120

[20] Ashley Karr. 2015. UX Research vs. UX Design. *Interactions* 22, 6 (Oct. 2015), 7–7. https://doi.org/10.1145/2834964

[21] Claire Kayacik, Sherol Chen, Signe Noerly, Jess Holbrook, Adam Roberts, and Douglas Eck. 2019. Identifying the Intersections: User Experience + Research Scientist Collaboration in a Generative Machine Learning Interface. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, Article CS09, 8 pages. https://doi.org/10.1145/3290607.3299059

[22] Eugene Kharitonov, Alexey Drutsa, and Pavel Serdyukov. 2017. Learning Sensitive Combinations of A/B Test Metrics. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (Cambridge, United Kingdom) *(WSDM '17)*. ACM, New York, NY, USA, 651–659. https://doi.org/10.1145/3018661.3018708

[23] Klaus Krippendorff. 2004. *Content Analysis: An Introduction to Its Methodology* (second ed.). Sage.

[24] Mike Kuniavsky, Elizabeth Churchill, and Molly Wright Steenson. 2017. *The 2017 AAAI Spring Symposium Series Technical Reports: Designing the User Experience of Machine Learning Systems*. Technical Report. Technical Report SS-17-04. Palo Alto, California.

[25] Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold P.O.S. Vermeeren, and Joke Kort. 2009. Understanding, Scoping and Defining User Experience: A Survey Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. ACM, New York, NY, USA, 719–728. https://doi.org/10.1145/1518701.1518813

[26] Effie Lai-Chong Law, Paul van Schaik, and Virpi Roto. 2014. Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies* 72, 6 (2014), 526–541.

[27] Alexandros Liapis, Christos Katsanos, Dimitris Sotiropoulos, Michalis Xenos, and Nikos Karousos. 2015. Subjective Assessment of Stress in HCI: A Study of the Valence-Arousal Scale Using Skin Conductance. In *Proceedings of the 11th*

Biannual Conference on Italian SIGCHI Chapter (Rome, Italy) *(CHItaly 2015)*. ACM, New York, NY, USA, 174–177. https://doi.org/10.1145/2808435.2808450
[28] Martin Lindvall, Jesper Molin, and Jonas Löwgren. 2018. From Machine Learning to Machine Teaching: The Importance of UX. *Interactions* 25, 6 (Oct. 2018), 52–57. https://doi.org/10.1145/3282860
[29] Megh Marathe and Kentaro Toyama. 2018. Semi-Automated Coding for Qualitative Research: A User-Centered Inquiry and Initial Prototypes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 348, 12 pages. https://doi.org/10.1145/3173574.3173922
[30] Marília Soares Mendes and Elizabeth Sucupira Furtado. 2017. UUX-Posts: A Tool for Extracting and Classifying Postings Related to the Use of a System. In *Proceedings of the 8th Latin American Conference on Human-Computer Interaction* (Antigua Guatemala, Guatemala) *(CLIHC '17)*. ACM, New York, NY, USA, Article 2, 8 pages. https://doi.org/10.1145/3151470.3151471
[31] Walter T. Nakamura, Elaine H. T. de Oliveira, and Tayana Conte. 2019. Negative Emotions, Positive Experience: What Are We Doing Wrong When Evaluating the UX?. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, Article LBW0281, 6 pages. https://doi.org/10.1145/3290607.3313000
[32] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. 2018. A Bermuda Triangle?: A Review of Method Application and Triangulation in User Experience Evaluation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 461, 16 pages. https://doi.org/10.1145/3173574.3174035
[33] B Joseph Pine, Joseph Pine, and James H Gilmore. 1999. *The experience economy: work is theatre & every business a stage*. Harvard Business Press.
[34] Christian Robert. 2014. Machine learning, a probabilistic perspective.
[35] V Roto, E Law, APOS Vermeeren, and J Hoonhout. 2011. U white paper-Bringing clarity to the concept of user experience. In *Outcome of Dagstuhl Seminar on Demarcating User Experience, Germany. http://allaboutux. org/uxwhitepaper*.
[36] Samaneh Soleimani and Effie Lai-Chong Law. 2017. What Can Self-Reports and Acoustic Data Analyses on Emotions Tell Us?. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. ACM, New York, NY, USA, 489–501. https://doi.org/10.1145/3064663.3064770
[37] Jacopo Staiano, María Menéndez, Alberto Battocchi, Antonella De Angeli, and Nicu Sebe. 2012. UX_Mate: From Facial Expressions to UX Evaluation. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. ACM, New York, NY, USA, 741–750. https://doi.org/10.1145/2317956.2318068
[38] Anselm Strauss and Juliet Corbin. 1990. *Basics of qualitative research*. Sage publications.
[39] The European Parliament and Council. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union* (2016).
[40] Alexandre N. Tuch, Rune Trusell, and Kasper Hornbæk. 2013. Analyzing Users' Narratives to Understand Experience with Interactive Products. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 2079–2088. https://doi.org/10.1145/2470654.2481285
[41] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 601.
[42] Terry Winograd. 2006. Shifting Viewpoints: Artificial Intelligence and Human–computer Interaction. *Artif. Intell.* 170, 18 (Dec. 2006), 1256–1258. https://doi.org/10.1016/j.artint.2006.10.011
[43] Qian Yang. 2017. The Role of Design in Creating Machine-Learning-Enhanced User Experience. In *2017 AAAI Spring Symposium Series*.
[44] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 130, 11 pages. https://doi.org/10.1145/3173574.3173704
[45] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating How Experienced UX Designers Effectively Work with Machine Learning. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. ACM, New York, NY, USA, 585–596. https://doi.org/10.1145/3196709.3196730
[46] Q. Yang, A. Steinfeld, C. Rosé, and J. Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020).
[47] Qian Yang, John Zimmerman, Aaron Steinfeld, and Anthony Tomasic. 2016. Planning Adaptive Mobile Experiences When Wireframing. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) *(DIS '16)*. ACM, New York, NY, USA, 565–576. https://doi.org/10.1145/2901790.2901858
[48] Xiang Zhang, Hans-Frederick Brown, and Anil Shankar. 2016. Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 5350–5359. https://doi.org/10.1145/2858036.2858523