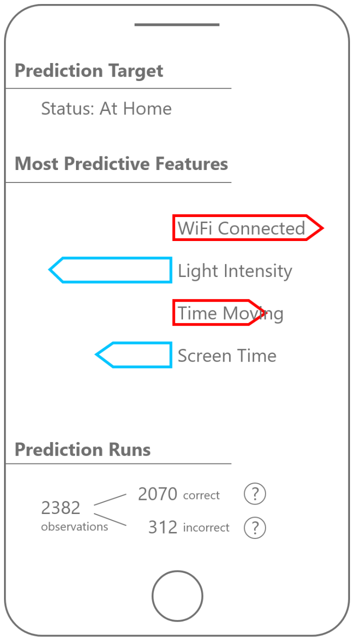

Figure 1: In a Mobile Sensing research app a locally trained personalized model is explained to the user, based on their own live sensing data. The user is included in the process of data collection and model development, and made aware of the hidden information that can be revealed from digital footprint data. Here: Unconspicuous WiFi status data is highly predictive of the user being at home.

Interactive End-User Machine Learning to Boost Explainability and Transparency of Digital Footprint Data

Florian Bemmann
LMU Munich
florian.bemmann@ifi.lmu.de

Daniel Buschek
Research Group HCI + AI, Department of Computer Science, University of Bayreuth
Bayreuth, Germany
daniel.buschek@uni-bayreuth.de

Heinrich Hussmann
LMU Munich
hussmann@ifi.lmu.de

HCXAI @ CHI 2021, May 08–13, 2021, Yokohama, Japan
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
Data collecting applications today only inform users about what data is collected directly, but not about what can be inferred from it. However, awareness of potential inferences is important from a data privacy perspective, especially as inferred information has been shown to be applicable for unethical applications as well. We propose interactive user involvement in model building: *Participatory Model Design* lets users interactively investigate what happens to their data, to convey which further information could be inferred. To operationalize such interactive explainability in practice, we created a prototype that integrates interactive personalized model training into a behaviour logging app for mobile sensing research. With our prototype we hope to spark discussions and further work towards strong direct user involvement in data collection and inference, to increase data privacy in the age of big data, and to facilitate explainability and transparency of downstream prediction systems.

Author Keywords
explainable artificial intelligence, data privacy, big data, machine learning

CCS Concepts
•**Human-centered computing** → *Interaction design theory, concepts and paradigms;* •**Computing methodologies** → **Philosophical/theoretical foundations of artificial in-**

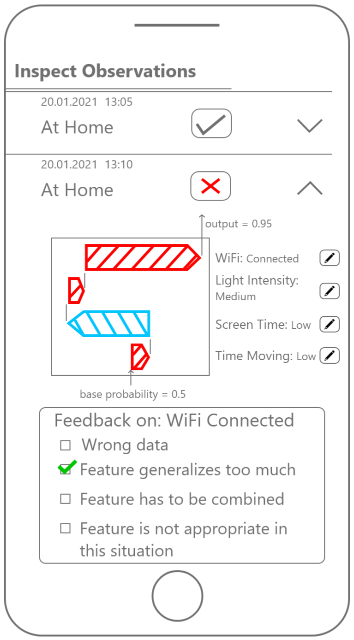

Figure 2: Interactively giving feedback on a locally trained personalized predictor: In the middle, a local explanation gives feedback from the system to the user, explaining which mobile sensing features contributed to the model's decision. At the bottom, the user gives feedback to the system, communicating why the model falsely concluded that the user is at home. Here: User states that the feature *WiFi connected* is too general.

telligence; •Security and privacy → *Human and societal aspects of security and privacy;*

Introduction

Systems that collect data about their users have become ubiquitous. Smartphone apps collect data about smartphone usage, the user's context and environment [11], web applications widely track how the user uses them, also including their personal preferences within the application's domain [4]. Such digital footprint data can fuel powerful artificial intelligence systems e.g. to predict future behaviour or characteristics of the user [6].

The purpose of predictions ranges from content personalization and recommendation [5], over adaptive user interfaces [15] to research purposes, e.g. in the fields of psychology [16].

With power comes responsibility: Tracking huge amounts of personal data demands for a good privacy protection concept to reach real transparency. The usual approach has been to inform the user about *who* collects and processes *what* data. However, with the possibilities which big data and psychometric modeling enable, that is not sufficient anymore. People should also be informed about *how* data is being used [9]. Therefore the full process of AI systems should be transparent [14].

Digital footprint data contains a huge amount of (concealed) information about their users. Using methods of psychometrics and psychological targeting [12], the data can be exploited for unethical purposes. Targeted advertisements based on digital footprint data can influence societies and poses huge challenges to our democracies [7].

Explaining through Interactivity for better Transparency

If users would be more aware of and understand how their data is leveraged for psychometric modeling, they may less likely be susceptible to content targeted based on their personality. By bringing explainability to the process of digital footprint model building, we think one can increase transparency and an understanding about how psychometrics work.

To do this we introduce participatory design and interactive explainable AI to data-collecting mobile apps. Instead of just being observed, users should be included in the full process of data collection and model building. Beyond showing what data is collected, it should be explained which features are extracted and what they are expressing, why the feature selection decides for certain features, and how a model can learn to predict a target variable from these inputs. Interactivity is therefore well suited, but unfortunately not very prominent in intelligent systems [1]. Research on intelligent systems calls for enabling rich feedback from the user towards the system [10], and interactive machine learning has demonstrated positive effects on learning [2] and explainability [8].

The concept of interactive user-involvement as explanations in XAI is contextually transferable and could be established as a general XAI technique. We envision more transparent and thus responsible intelligent systems, by letting users participate in and interact with data collection and model building.

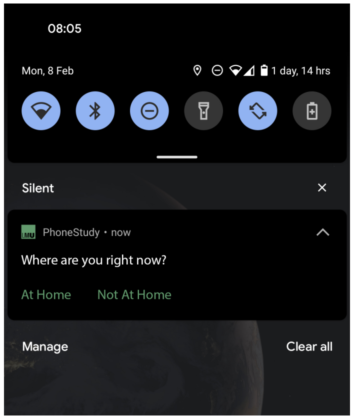

Figure 3: To collect ground-truth data for mobile model building in our prototype, the user was asked to indicate when leaving from / returning to home via a permanent notification

Proof of Concept: Interactive Model Building Demonstrates the Hidden Information in Digital Footprint Data

To operationalize interactive explainability in practice, we incorporate it into a Mobile Sensing smartphone app. Our basis is the PhoneStudy app[1], which collects passive sensing data and self-reports in the wild to fuel offline model building for psychological research [16].

In a working prototype we brought the model building to the client device, allowing the user to train a personalized model on-device. Features are extracted locally from user's live sensing data, and self-reported data is used as prediction target.

In a first experiment we use the binary variable *being at home or not at home* as prediction target and basic device status features as input data. Although this is a trivial example, it demonstrates our concept of mobile model building with live sensing data to show to users how "hidden" information in digital footprint data might be revealed: The model fitted nearly perfectly on features on the smartphone's WiFi status.

Discussion

Can interactivity boost explainability?
We argue for interactivity – going beyond passive explanations. The literature supports this idea: For example, trying configurations of learning systems and observing effects is desired by users of ML systems and could result in a better understanding [8] and long-term learning effects. The effects of absence and presence of specific features on the model performance could convey their value to the user. Principles of correctability and rich feedback [10] should be incorporated in ML systems, and be used beyond improving models. To further inform and evaluate this approach, it should be studied how such interactivity and user-to-system feedback affects the understanding of a system: Which insights can be conveyed easier, and which cannot?

Interactive ML for Data Transparency
Transparency is limited in current systems that collect data for model building: Beyond explaining the user *who* collects *what* data, systems should convey *how* data can be used and what can be *inferred* [9] to reach real transparency. To implement this, we suggest to include the user in the model building process. We want to study whether interactively trying out inference techniques with their own data, thus experiencing which kinds of predictions can work and which are more difficult, support real transparency.

Participatory Model Design
In our concept the user is included in the model building process, rather than just observed [1]. To reach "Participatory Model Design", inspired by Participatory (Product) Design, new workflows for research working with user data should be studied: Can users be involved in the big challenges, e.g. feature design and selection? Users could create features they think are predictive for them, try them out in a local model, refine their features. Finally the researchers collect only a set of individual models and feature descriptions from their study participants.

Interactivity for Personalized vs. Universal Models
Personalized models can demonstrate predictions about intra-user variables, such as some status of the user (e.g. indoor/outdoor, mood, stress, ...). To showcase models that compare users (inter-user variables) among certain characteristics (e.g. personality) a universal model is needed. While for a personalized model the full process of training, evaluation and prediction can be demonstrated live, for uni-

[1] https://osf.io/ut42y/, last accessed 7th February 2021

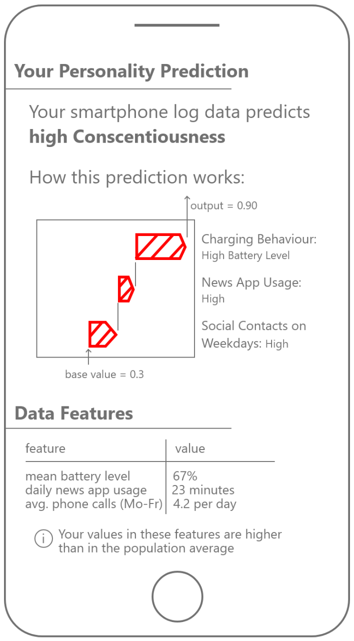

Figure 4: Interactivity on personalized predictors is only suitable for intra-user variables. To explain inter-user variables like personality, we propose to deploy a pre-trained predictor on the client device and locally run and explain predictions using the data collected by the user's device.

versal models it is only possible to show predictions using a pre-trained model. Thus, different concepts of interactivity have to be designed for both types of models, and their effects on the user may have to be studied separately.

Make a Difference: Inference Potential and (Unethical) Applications
Users being aware of what can be inferred from their digital footprints is an important first step. However, to make a difference with this work we encourage to think beyond: Building on the outlined interactive explainable mobile sensing app, it should be studied how it can further be used to "vaccine" people against (unethical) applications of personal data collection: McGuire's Inoculation Theory [13] proposes weakened pre-exposure to protect against persuasion. With *The Bad News Game*[2] the application of Inoculation Theory has shown a reduction of susceptibility in the domain of online misinformation [3]. Similar concepts seem promising to the domain of targeted content as well. Can we demonstrate the suggestibility of content that is targeted with personal data, to empower users to unmask and reluct against such in the future?

Motivating User Engagement
Our concept is not targeted to a specific user-group, instead any data generating smartphone user should be encouraged to use it. While a short term usage could be motivated by Gamification techniques (e.g. little challenges or tasks), we assume that long-term use would require more sophisticated application concepts. Furthermore it should be discussed whether long-term use is even needed to yield the desired understanding.

[2] https://www.getbadnews.com, last accessed 22nd March 2021

Acknowledgements
This project is partly funded by the Bavarian State Ministry of Science and the Arts and coordinated by the Bavarian Research Institute for Digital Transformation (bidt).

REFERENCES
[1] Nitin Agrawal, Reuben Binns, Max Van Kleek, Kim Laine, and Nigel Shadbolt. 2021. Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation. *arXiv:2101.08048 [cs]* (Jan. 2021). DOI: `http://dx.doi.org/10.1145/3411764.3445677` arXiv: 2101.08048.

[2] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. 2014. Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine* 35, 4 (Dec. 2014), 105–120. DOI: `http://dx.doi.org/10.1609/aimag.v35i4.2513` Number: 4.

[3] Melisa Basol, Jon Roozenbeek, and Sander van der Linden. 2020. Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of cognition* 3, 1 (2020).

[4] Tomasz Bujlow, Valentín Carela-Español, Josep Sole-Pareta, and Pere Barlet-Ros. 2017. A survey on web tracking: Mechanisms, implications, and defenses. *Proc. IEEE* 105, 8 (2017), 1476–1510.

[5] Hong Cao and Miao Lin. 2017. Mining smartphone data for app usage prediction and recommendations: A survey. *Pervasive and Mobile Computing* 37 (June 2017), 1–22. DOI: `http://dx.doi.org/10.1016/j.pmcj.2017.01.007`

[6] Trinh Minh Tri Do and Daniel Gatica-Perez. 2014. Where and what: Using smartphones to predict next

locations and applications in daily life. *Pervasive and Mobile Computing* 12 (2014), 79–91.

[7] Gibbs A Doward J. 2017. Did Cambridge Analytica Influence the Brexit Vote and the US Election? (2017). https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump

[8] Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, and Heinrich Hussmann. 2019. When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications. (2019), 11.

[9] Gabriella M Harari. 2020. A process-oriented approach to respecting privacy in the context of mobile phone tracking. *Current Opinion in Psychology* 31 (Feb. 2020), 141–147. DOI: http://dx.doi.org/10.1016/j.copsyc.2019.09.007

[10] Todd Kulesza, Margaret Burnett, Weng-Keen Wong, and Simone Stumpf. 2015. Principles of Explanatory Debugging to Personalize Interactive Machine Learning. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, Atlanta Georgia USA, 126–137. DOI: http://dx.doi.org/10.1145/2678025.2701399

[11] Francisco Laport-López, Emilio Serrano, Javier Bajo, and Andrew T. Campbell. 2020. A review of mobile sensing systems, applications, and opportunities. *Knowledge and Information Systems* 62, 1 (Jan. 2020), 145–174. DOI: http://dx.doi.org/10.1007/s10115-019-01346-1

[12] S. C. Matz, M. Kosinski, G. Nave, and D. J. Stillwell. 2017. Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences* 114, 48 (2017), 12714–12719. DOI: http://dx.doi.org/10.1073/pnas.1710966114 Publisher: National Academy of Sciences _eprint: https://www.pnas.org/content/114/48/12714.full.pdf.

[13] William J McGuire. 1964. Inducing resistance to persuasion. Some contemporary approaches. 1 (1964), 191–229.

[14] Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission. 2019. *Ethics Guidelines for Trustworthy AI*. European Union. https://ec.europa.eu/futurium/en/ai-alliance-consultation

[15] Choonsung Shin, Jin-Hyuk Hong, and Anind K. Dey. 2012. Understanding and Prediction of Mobile Application Usage for Smart Phones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 173–182. DOI: http://dx.doi.org/10.1145/2370216.2370243

[16] Clemens Stachl, Quay Au, Ramona Schoedel, Samuel D Gosling, Gabriella M Harari, Daniel Buschek, Sarah Theres Völkel, Tobias Schuwerk, Michelle Oldemeier, Theresa Ullmann, and others. 2020. Predicting personality from patterns of behavior collected with smartphones. *Proceedings of the National Academy of Sciences* 117, 30 (2020), 17680–17687.