User Mental Models of AI-based Personal Informatics

Corina Sas

Lancaster University, UK, c.sas@lancaster.ac.uk

The rather separate growth of two strands of work on personal informatics and AI technologies has prompted increased interest in AI-based personal information technologies. Such work however faces the challenges of user limited mental models of technologies in general and AI ones in particular, and of their ethical concerns. To address such challenges, the paper proposes to build on HCI work on user mental models, sociotechnical infrastructures and ethical design. The paper concludes with a reflection on our findings in these areas and articulates novel directions to address the identified challenges in order to design more ethical AI-based personal informatics systems.

CCS CONCEPTS • Human-centered computing ~Human computer interaction (HCI)

Additional Keywords and Phrases: Mental models, Sociotechnical infrastructures, Ethics, Value-creation.

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

The third wave HCI and its focus on everyday life has led to increased interest in personal informatics technologies, explored by HCI scholars through mobile, tangible or wearable systems, aimed to provoke, delight, motivate or support change. While research prototypes of personal informatics technologies lead to "small data" such as those collected usually through limited interactions in lab-based evaluation studies, the working prototypes deployed and evaluated through longitudinal studies in the wild, can lead to "big data" sets. The same is the case for commercial personal information systems from mobile apps [8][22] to online platforms such as SilverCloud providing cCBT programs [35].

In parallel, we have seen a growth of AI technologies which is now prompting increased HCI interest in how to extend the body of work on personal informatics towards designing AI-based personal informatics systems. However, despite the emerging work on explainable AI [3][11], we know little of how AI-based personal informatics systems can be better understood, and designed for to account also for their ethical challenges.

In contrast to traditional interactive technologies, AI-based technologies present additional challenges for the development of users' mental models [31], originated in users' limited understanding of how AI-based technologies more broadly and in particular personal informatics ones work. The latter pertain to users' personal data and their ethical implications [18] including awareness of the flow of data as well as its value generation. The current

discourse on AI-based technologies and their impact on users' privacy and trust [24] reflects these challenges. However, while most work on explainable AI has focused on making transparent the process through which specific user recommendations or decisions are inferred through deep learning, there have been limited efforts to explore explainable AI from a human-centered approach [1][11][32][35], and even less so for understanding the flow of data and its value generation. Indeed, most traditional user models have focused on technologies leveraging merely personal data [29]. In contrast, the integration of personal data into big data has received limited attention. User mental models reflecting this integration will not be trivial to develop, as they will most likely require an infrastructure perspective including multiple stakeholders and data sources. In addition, the value of personal data for AI-based technologies positions users in a new role, beyond the mere consumers but also as key stakeholders without whose data, such technologies would not be possible. In this position paper I argue for two promising directions to address such challenges by leveraging HCI research in user mental models, and the emerging work on designing for sociotechnical infrastructures. The paper frames AI-based personal informatics technologies as sociotechnical ones where data is gathered across distributed users and their personal informatics technologies.

2 USER'S MENTAL MODELS: SOCIOTECHNICAL INFRASTRUCTURES

HCI research on mental models has consistently shown their importance for helping users to better complete a broad range of tasks [25][26][27][30]. Such work has also highlighted the distinction between functional or superficial mental models developed by novice users supporting them to use a system, and the structural models developed by expert users and reflecting deeper understanding of systems' inner working [12][29][33]. This distinction emphasizes the challenges of novice or less technical users to develop mental models of everyday technologies.

While most work has looked at personal informatics technologies [1][2][19][34], less research has explored mental models [9] of large scale sociotechnical infrastructures [35] that involve the integration of actors, materials and information [15][28][35]. Sociotechnical infrastructures such as those underpinning AI-based technologies, be them personal informatics systems in particular, or blockchain [15][16][28][29] and IoT technologies [30] more broadly, have been mostly explored from technological lens. In this respect, HCI work for understanding user mental models of such complex, distributed technologies has started to emerge. However, exploration of novel tools supporting the understanding and development of these models is rather limited.

An illustration of such tools for supporting user mental models of blockchain sociotechnical infrastructures is BlocKit [17]. While traditional tools for representing mental models tend to be in text or diagrammatic form [13], BlocKit was designed to offer a tangible representation of blockchain infrastructure. Building on material-centred design and embodied cognition, it metaphorically represents key concepts such as bitcoins, wallets, passwords, private and public keys, transactions through physical objects made of clay, paper or transparent containers. The affordances of these objects allow users to manipulate them in ways that support the understanding of these concepts. An evaluation study with blockchain experts has shown the value of BlocKit to support high level of engagement while communication and designing for blockchain infrastructure. More importantly, BlocKit has also strengthened experts' structural mental models of the inner working of the blockchain infrastructure [17].

Another example of sociotechnical infrastructure is that explored in developing contexts for making of, and selling cultural heritage artifacts such as songket fabric. Such infrastructure involves diverse and geographically distributed social actors from weavers, designers and merchants to customers who despite different levels of digital literacy, successfully engage with technologies such as smartphones, tablets, computers, drawing software,

Facebook, Instagram or WhatsApp, and online payment services to ensure their needed flow of information and materials [36]. Such work revealed the importance of understanding and designing for this rich social layer of infrastructure, and the sensitivity needed to account for the interdependency of its actors engaged in "mutually advantageous exploitative relationships". The findings also highlighted hybrid artifacts capturing the songket from its design in digital and physical form, to its materialization in fabric, and how the novelty of such design ensures value creation within the songket's supply chain [26].

3 REFLECTION

I now reflect on how these findings can offer new insights into how to address the above challenges of designing more ethical AI-based personal informatics systems. First, is the importance of extending explainable AI through novel tools for materializing users' mental models of such technologies, supporting users' understanding of, and engagement with them, but also for questioning their ethical underpinnings. Here, we can draw inspiration from design cards which help articulate abstract concepts of technology acceptance theories [20][21] or BlocKit leveraging embodied cognition metaphors to materialize high level concepts of blockchain infrastructure [17].

Second, our framing of AI-based personal informatics technologies as sociotechnical ones where data is gathered across distributed users and their diverse technologies is important. As shown by findings on songket's supply chain [26], the social layer of infrastructure matters as do the digital-physical artifacts supporting value generation. For the former, we argue for the importance of going beyond the mere focus on users of AI-based personal informatics, but to consider the large social layer including other users, developers and third party entities with whom one shares their data. Full transparency in this respect will support building trust and long term adoption of such technologies.

Third, such materialization should capture not only the flow of data but also the flow of data as asset and its generated value, alongside with the novel business models allowing users to become co-beneficiary of such value, beyond the mere personalization. Indeed, tthrough their continuous use and the provision of their data, users become ongoing co-creators of their AI-based technologies. Those benefiting from these are users themselves, other users sharing similar characteristics, as well as developers of these technologies, albeit only the latter benefit directly from monetizing such data [23]. This poses interesting questions from an ethical perspective [25], given the potential of AI-technologies for the 4th industrial revolution and its anticipated impact on labor market [14].

Finally, with regard to ethics, the significant impact of AI-based technologies in our everyday lives has led to a growing number of ethical concerns and their pressing demands to be addressed. For specific domains such as digital wellbeing [4] or affective health [5][6][9][10], AI-based such technologies feeding into big data raise even stronger ethical concerns, given its highly private, sensitive content , or biodata which may further challenge users' anonymity [25]. This begs the question of how to balance the benefits of big data for personalization with such ethical challenges. The two points above outline novel approaches to answer this question by innovative business models that allow sharing the value being created with users themselves, and materializing the flow of such data and value to support users' understanding of such technologies and their informed decisions on data use.

ACKNOWLEDGMENTS

This work was supported by the Innovative Training Network "Marie Curie Actions" funded by the H2020 People Programme (GA 722022) entitled AffecTech: Personal Technologies for Affective Health.

REFERENCES

[1] Alan Dix, and Corina Sas. 2008. Public displays and private devices: A design space analysis. In *Workshop on designing and evaluating mobile phone-based interaction with public displays*. CHI2008, Florence, 5 April 2008

[2] Alan Dix, and Corina Sas. 2010. Mobile personal devices meet situated public displays: Synergies and opportunities. *International Journal of Ubiquitous Computing* 1, no. 1 (2010): 11-28.

[3] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y Lim, and Mohan Kankanhalli. 2018. Trends and trajectories for explainable, accountable and intelligible systems: An HCI research agenda. In *Proc. CHI'18,* 1–18.

[4] Sultan Almoallim, Corina Sas 2022. Functionalities Review of Digital Wellbeing Apps: Towards Research-informed Design Implications for Interventions Limiting Smartphone Use. *JMIR Formative Research.*

[5] Dionne Bowie-DaBreo, Sandra, Sünram-Lea, Corina Sas, Hearther Iles-Smith, H. 2020. Evaluation of depression app store treatment descriptions and alignment with clinical guidance: Systematic search and content analysis. *JMIR Formative Research,* 4(11), 18 pages.

[6] Dionne Bowie-DaBreo, Corina Sas, Heather Iles-Smith, Sandra Sunram-Lea. 2022. User perspectives and ethical experiences of apps for depression: A qualitative analysis of user reviews. In *Conference on Human Factors in Computing Systems (CHI'22).*

[7] Claudia Daudén Roquet and Corina Sas. 2020. Body Matters: Exploration of the Human Body as a Resource for the Design of Technologies for Meditation. *Proceedings of Designing Interactive Systems Conference*. ACM, 533–546.

[8] Dauden Roquet, C., & Sas, C. 2018. Evaluating mindfulness meditation apps. In CHI *Extended Abstracts (CHI EA '18*) LBW575, 6 pages.

[9] Claudia Daudén Roquet and Corina Sas. 2021. Interoceptive Interaction: An Embodied Metaphor Inspired Approach to Designing for Meditation. In *Conference on Human Factors in Computing Systems* (*CHI '21*). ACM, 1–17.

[10] Claudia Daudén Roquet, Corina Sas & Dominic Potts (2021) Exploring Anima: a brain–computer interface for peripheral materialization of mindfulness states during mandala coloring, *Human–Computer Interaction*, 37(4), 41 pages.

[11] Upol Ehsan and Mark O Riedl. 2020. Human-centered Explainable AI: Towards a Reflective Sociotechnical Approach. arXiv preprint arXiv:2002.01092

[12] Saul Greenberg and Chester Fitchett. 2001. Phidgets: Easy Development of Physical Interfaces through Physical Widgets. In *Proc. UIST'01* 209–18

[13] Mary Hegarty and Marcel Adam Just. 1993. Constructing Mental Models of Machines from Text and Diagrams. In *Journal of Memory and Language.* NY, 32, 6: 717-742.

[14] Yuval Noah, Harari. 2017. Reboot for the AI revolution. *Nature News* 550, 7676 (2017): 324.

[15] Irni Eliana Khairuddin and Corina Sas. 2019. An Exploration of Bitcoin Mining Practices: Miners' Trust Challenges and Motivations. In *Proc. CHI'19,* p.1-13.

[16] Irni Eliana Khairuddin and Corina Sas. 2016. Exploring Motivations for Bitcoin Technology Usage. In *Extended Abstracts* CHI'16, 2872-2878.

[17] Irni Eliana Khairuddin, Corina Sas, and Chris Speed. 2020. BlocKit: A physical kit for materializing and designing for Blockchain infrastructure. In *Proc. DIS'19*, 1449-1462. 2019.

[18] Will Knight 2017. The dark secret at the heart of al. *Technology Review* 120, no. 3 (2017): 54-61.

[19] Christian Kray, Keith Cheverst, Dan Fitton, Corina Sas, John Patterson, Mark Rouncefield, Christoph Stahl. 2006. Sharing control of dispersed situated displays between and residential users. In *Proc. Conference on Human-Computer interaction with Mobile Devices and Services*,. 61-68.

[20] Camille Nadal, Shane McCully, Kevin Doherty, Corina Sas and Gavin Doherty. 2022. The TAC toolkit: Supporting the design for user acceptance of health technologies from a macro-temporal perspective. In *Proc. Conference on Human Factors in Computing Systems (CHI'22).*

[21] Camille Nadal, Corina, Sas, Gavin Doherty. 2020. Technology acceptance in mHealth: a scoping review of definitions, models and measurement. *Journal of Medical Internet Research,* 22(7), 17 pages.

[22] Chengcheng Qu, Corina Sas, Claudia Dauden Roquet, Gavin Doherty 2020. Functionality of top-rated mobile apps for depression: Systematic search and evaluation. *JMIR mental health*, *7*(1), 13 pages.

[23] Jeffrey F Rayport and John J Sviokla. Exploiting the virtual value chain. *Harvard business review*, 73(6):75, 1995.

[24] Alberto Romele, Francesco Gallino, Camilla Emmenegger, and Daniele Gorgone. Panopticism is not enough: Social media as technologies of voluntary servitude. *Surveillance & Society*, 15(2):204–221, 2017.

[25] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. HCI and affective health: Taking stock of a decade of studies and charting future research directions. In *Proc. CHI'19*, page 1–17, ACM.

[26] Corina Sas, R. Reilly and Gregory M. P. O'Here. 2003. A Connectionist Model of Spatial Knowledge Acquisition in a Virtual Environment. In *Proc. User Modeling, Information Retrieval and Machine Learning.*

[27] Corina Sas. 2004. Individual Differences in Virtual Environments. In LNCS, Vol 3038. Springer.

[28] Corina Sas, Irni Eliana Khairuddin. 2015. Exploring Trust in Bitcoin Technology: A Framework for HCI Research. In *Proc. OzCHI* '15, 338- 342.

[29] Corina Sas and Irni Eliana Khairuddin. 2017. Design for Trust: An Exploration of the Challenges and Opportunities of Bitcoin Users. In *Proc. CHI' 17*, 6499-6510.

[30] Corina Sas,Carman Neustaedter. 2017. Exploring DIY Practices of Complex Home Technologies. In *ACM Trans. Comput.-Hum. Interact*., 24, 2.

[31] Nancy Staggers and A.F Norcio. 1993. Mental Models: Concepts for Human-computer Interaction Research. *International Journal of Man-Machine Studies*, 38: 587-605.

[32] Corina Sas, Kristina Höök, Gavin Doherty, Pedro Sanches, Tim Leufkens, and Joyce Westerink. 2020. Mental Wellbeing: Future Agenda Drawing from Design, HCI and Big Data. In *DIS' 20 Companion*. 425–428.

[33] Muhammad Umair, Corina Sas, and Miquel Alfaras. 2020. ThermoPixels: Toolkit for Personalizing Arousal-based Interfaces through Hybrid Crafting. In *Proc. DIS '20*, 1017-1032.

[34] Muhammad Umair, Corina Sas, and Muhammad Hamza Latif. 2019. Towards affective chronometry: Exploring smart materials and actuators for real-time representation of changes in arousal. In *Proc. DIS'19*, 1479–1494

[35] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems. *ACM Trans. Comput.-Hum. Interact.* 27, 5, Article 34, 53 pages.

[36] Min Zhang, Corina Sas, Zoe Lambert and Masitah Ahmad. 2019. Designing for the Infrastructure of the Supply Chain of Malay Handwoven Songket in Terengganu. In *Proc. CHI'19*, 1-15.