Human-Data Interaction through Design

An explorative step from theory to practice using Design as a vehicle

NEELIMA SAILAJA

Horizon Digital Economy Hub, University of Nottingham, Nottingham, United Kingdom.

JOSEPH LINDLEY

Imagination Lancaster, University of Lancaster, Lancaster, United Kingdom.

LACHLAN URQUHART

School of Law, University of Edinburgh, Edinburgh, United Kingdom.

DEREK MCAULEY

Horizon Digital Economy Hub, University of Nottingham, Nottingham, United Kingdom.

IAN FORRESTER

BBC R&D, BBC, Salford, United Kingdom.

The increasing use of personal data and AI in everyday technologies has resulted in the amplification of complex and intertwined socio-technical challenges. These, often exemplified by data abuse, breaches, and exploitation, must be alleviated to support sustainable, resilient and human-centred data economies and positive global innovation. Here, we turn towards Human-Data Interaction, an interdisciplinary branch of research, inspired by HCI, that brings together diverse siloed perspectives to present three holistic response principles: data legibility, negotiability and agency. But, the emergent nature of this field calls for refinement of these theoretical tenets to help them translate into practical and tangible responses that are embedded in the technologies we create. We propose this workshop as a foundational step towards this agenda by opening these principles to the CHI community to encourage critique and dialogue about the strengths, weaknesses, value and opportunities of incorporating HDI into the design and evaluation of technology. The outcomes of this workshop, by engaging with HDI through Design, will form the basis for the next stages of research within HDI by contributing to foundational texts within academia and implementing HDI-infused systems within industry.

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

Additional Keywords and Phrases: Human Data Interaction, Personal Data, Artificial Intelligence, Design Research

1 BACKGROUND AND RATIONALE

The emergence of consumer IoT devices coupled with the 'rebirth' of AI is seeing innovation in data-driven technologies across domains ranging from finance to retail to media. This phenomenon results in a significant increase in the everyday users' 'data interactions' wherein users engage both actively and passively in processes pertaining to their personal data. This increased collection and use of personal data has been shown to result in a number of socio-technical implications [12], particularly owing to highly publicised data breach [1,2] and data abuse [6] instances. Here, as innovation frequently centres on leveraging personal data these challenges become proportionally more pronounced, calling for immediate response.

Both industry and academia are actively seeking interventions that alleviate these challenges, which have been studied across different disciplines. In order to move towards effective solutions to these overarching societal implications there is the need to bring together disparate perspectives and start addressing these challenges comprehensively, in a holistic manner. The Human Data Interaction (HDI) movement aims to do this and takes direct inspiration from the dynamic and adaptive nature of the broader HCI community. HDI takes an interdisciplinary approach to engage with contemporary socio-technical challenges through an amalgamation of prior insights, methods and perspectives from multiple satellite disciplines such as sociology, psychology, law and computing. Adopting this hybrid approach, the fledgling HDI movement proposed three key principles - data legibility, negotiability and agency - which have found traction in recent years' HCI conferences [10,13].

As the HDI movement gains momentum and matures there is a need to clarify and stratify the constituent parts of the HDI research programme in order to establish a productive HDI paradigm. New perspectives in HCI often emerge at the confluence of fundamental/ theoretical research contributions, and practical/ applied research contributions. Paradigms emerge through interplay between these strata; theory informs practice and examples of practical work constitute a body of evidence to support, critique, and progress theory. Here, the need and urgency of this translation of theory to tangible responses that effectively alleviate said wider socio-technical challenges within technologies is explicitly highlighted in the most recent U.K. Parliament Sub-Committee Hearing on Online Harms and Ethics of Data [7]. The aim of this workshop is to initiate a research programme and agenda for HDI that supports this interplay and situates it against the broader HCI zeitgeist, thereby marking a key step towards realizing effective HDI-infused responses. In doing so the aim is to move towards a concrete and productive paradigm for HDI as it relates to SIGCHI as well as for the computing community more generally.

The workshop's programme will be situated in relation to our own varied experiences working with, and around, HDI. This includes contributing to HDI's foundational texts [11]; implementing HDI-infused systems in a global media corporation [5,12,13] and engaging with HDI challenges through Design Research [10,16]. In addition to the organisers' own experience we wish to draw on examples and proposals offered by workshop participants through submitted papers, proposals or provocations. By sharing and collating a diverse set of experiences working with HDI we will focus on what factors facilitate productive applications of HDI in practice (and conversely, what factors hinder applications of HDI or stifle the movement's progress). Specific questions we wish to address include:

1. Are the current HDI concepts still the right ones as prevalence of complex systems have emerged?
2. How do new human-AI interactions challenge the tenets, given the complexity and 'black boxing' of such systems?
3. How could/ should HDI's tenets be refined to promote impact?

The workshop will proceed through three stages. First, we will present the organisers' view of the state-of-the-art in HDI research to the workshop, giving attendees an opportunity to highlight the successes and failures of HDI implementation in these instances. Second, we will facilitate participants to consider practical design revisions and interventions that would effectively enable realisation of the tenets of HDI within sample technologies. Third, we will use the specific, contextualised design processes, decisions and revisions surfaced in the first two sections to produce a commentary on how HDI has accelerated or impeded the design and adoption of technology. The aim is to produce a concise agenda for the future of HDI research, mapping out the research questions which need to be addressed to support and further the goals of the HDI movement; successful and impactful implementation of data driven technologies.

There is a clear need for HDI. However, ensuring that HDI is current and coherent is a significant challenge. Moreover, transitioning from theory and a handful of examples, to widespread global impact requires sustained effort. Hence, the workshop aims to establish an agenda in order to refine, empower, and promote HDI.

2 ORGANISERS

2.1   Dr. Neelima Sailaja

Neelima Sailaja is a Research Fellow at the Horizon Digital Economy Hub at the University of Nottingham where she studies the socio-technical implications of personal data use in media experiences, particularly technological responses to these challenges. In doing so, she employs HDI, HCI, ethics and design, exploring ideas around personalisation, artificial intelligence, data privacy and trust. In this work, which has also recently received formal support from the HDI-network, she works closely with BBC teams working on next generation media experiences, studying and developing data-sensitive responses, work processes and best practice guidelines which informs industry [14] and policy [15] in this changing scape.

2.2   Dr. Joseph Lindley

Joseph Lindley is a Research Fellow at the Imagination Lancaster research lab where his work promotes the use of Design Research to explore and understand the social implications of emerging technologies. His recent research utilised Design Research in order to apply Human-Data Interaction principles to contemporary AI systems. He is a member of the ACM Future of Computing Academy, contributes to the IEEE Working Group for Responsible AI Licensing, and recently co-authored a report for the UK Arts and Humanities Research Council on AI and Data.

2.3   Dr. Lachlan Urquhart

Lachlan Urquhart is a Lecturer in Technology Law at the University of Edinburgh and Visiting Researcher at Horizon Digital Economy Hub, University of Nottingham. He has an interdisciplinary background in Computer Science (PhD) and Law (LL.B; LL.M) where his main research interests are human computer interaction, ubiquitous computing, data protection and cybersecurity. He has engaged in previous research looking at bridging the 'realisation' gap between laws like the GDPR and their practical inclusion within data driven technologies and seeks to explore this further through HDI.

2.4   Prof. Derek McAuley

Derek McAuley is a Professor of Digital Economy in the School of Computer Science at the University of Nottingham and Director of Horizon Digital Economy Research, an interdisciplinary research institute funded through the RCUK Digital Economy programme. He acted as Specialist Advisor to the House of Lords European Union Committee into online platforms, and Chief Innovation Officer during the founding of the Digital Catapult in the U.K.. Being one of the key contributors to the foundational theory behind HDI, he seeks to further this agenda by identifying methods of translating this concept to tangible responses within everyday technologies.

2.5   Ian Forrester

Ian Forrester is a Senior 'Fire-Starter' Producer at BBC R&D. His engagement with the implications of personal data use in media has been long and diverse. He had invested in several industry initiatives around this theme both within and outside the BBC including a video series showcasing the views of BBC staff members to the turn towards personal data use, implications of this shift and potential responses. He has, on several occasions brought together industry and academic research to produce provocations and solutions that could intervene in this area, which have been showcased at prestigious national and international venues. He would be leveraging this extensive knowledge base to provide solid examples of HDI inclusion within media technologies and extend the impact of this agenda, particularly within industry circles.

3 WEBSITE

A website would be hosted online for the purpose of community building and recruitment for the workshop. The website would present detailed information about the following :

3.1   The workshop theme

1. Human Data Interaction : Tenets and Theory

3.2   The motivations

1. Explication of the novel nature of HDI
2. The urgency and need for translating this concept to tangible responses within technologies

3.3   Aims and objectives

1. Present the domain of HDI to the wider CHI community
2. Allow opportunity for critique and commentary of this agenda
3. Initiating foundation for the next steps (in terms of research agenda) for this emergent field

3.4   Other previous events, activities and publications engaged in

1. Introduction to the HDI-network within the U.K. [8]
2. Examples of Academic Contributions [10,13]
3. Examples of Industry Engagements [3,5]

3.5   Workshop structure

1. As detailed in Section 5 of this proposal

3.6   Workshop participation

1. Details of the position paper and space for upload of this document
2. An alternative online form for expression of interest of participation

3.7   Future possibilities for engagement and collaboration enabled by the workshop

1. Links to current research around this agenda
2. A list of planned future activities with contact details for future collaborations

3.8   Contact Details of the organisers

This website would be used for recruitment and communication of updates regarding the workshop logistics. It would also be integrated with social media accounts of the organisers and network partners, for easy communication that supports further community building (URL to be assigned post proposal acceptance).

4   PRE-WORKSHOP PLANS

Recruitment for the workshop would be done through the website. It would allow participants to express interest in two forms: a traditional workshop 'position paper' or a more light-weight online form on the website in order to convey a proposal or provocation. The position paper would be in the ACM Master Article Submission Templates format and would be a maximum of 2 pages (excluding references). The submission should either describe an example of employing HDI in the context of research or innovation explaining why HDI was used, what it helped to achieve (or failed to achieve), and how other research/ innovations may benefit from using it in a similar way. Or it could offer a critique or comment on HDI's tenets, explain - using examples, theory or rhetoric - why these are the right (or the wrong) concepts and what their strengths or limitations are.

This dual approach is intended to allow those that would benefit from the process of more formally articulating their contribution (via a workshop paper) to attend, whilst also keeping the workshop open to participants who may prefer a less academic or briefer way to contribute. We are allowing for both due to HDI being an interdisciplinary area, with much interest from outside the academy, we believe that providing a more informal point of entry along with the possibility of writing a strong position paper, as necessary.

Participants will be selected based on their prior experience of HDI (or related areas of interest) and the quality/ relevance of their submission to the workshop theme. Applicants will be selected in order to compose a group of complementary participants whose submissions will stimulate a lively and productive discussion. We will endeavour to assemble a diverse (in terms of career stage, race, gender, location) group of participants.

The participants would be provided access to a Slack channel before the workshop event. This would be used as a means of dissemination of information and instructions to the participants, provide a platform for the participants to introduce themselves and get involved with the community *prior* to the event itself, it will provide a persistent means to share content *during* the workshop, and leveraged for future collaborations, report writing and special issue engagement *post* the workshop.

5 WORKSHOP STRUCTURE

The workshop is designed for a maximum of 20 participants, as a single day's 4 hour session (including two planned 15 minute breaks in between). While all organisers would be actively involved throughout the workshop, each session would be led by a different organiser, depending upon expertise and interest. We also intend to involve volunteers/ facilitation assistants from our research pool to further minimise the organiser-participant ratio and ensure maximum participant engagement, particularly in the co-design sessions.

5.1   Activity 1 : Orientation [30 minutes]

The workshop will begin with an orientation session that presents Human-Data Interaction, its motivations, aims, the three principles and the theoretical underpinnings that support their ability to respond to the socio-technical challenges of personal data leverage. This would be achieved through a presentation led by the workshop organisers followed by a warm-up activity that helps introduce the participants to each other and to the topic in hand. The session would also include time spent on gaining participant consent before the activities begin. The warm-up activity will engage the entire group as a whole and serve as an ice-breaker, it will utilise a simple facilitation technique based on participants sharing their perspectives in a sentence or less via the form "The best thing about HDI is …" and "The worst thing about HDI is …".

5.2   Activity 2 : State-of-the-art showcase [30 minutes]

In this session the organisers will present recent examples of inclusion of HDI principles within data driven technologies. This would include a series of videos that present work around experimental technologies that leverage substantial amounts of user data in order to implement Machine Learning enabled features. After each video, the organisers would highlight the HDI inclusions in these technologies, the processes that led to it and user response to them. The projects to be showcased here are the Living Room of the Future/ GoggleBox [9,13], AI Icons [10] and a BBC Cross Service Recommender. As this session progresses we will provide participants with appropriate solutions to capture their views on the strengths/ weaknesses of how HDI has been implemented in each case, these insights will feed into activity 3.

5.3   Break [15 minutes]

5.4   Activity 3 : Critique State-of-the-art [30 minutes]

This session entails a group activity for which the larger group would be divided into smaller groups of 3-4 people, for more engaged discussions and response. The participants would be encouraged, in their groups, to select one or more of the examples presented and critique the inclusion of HDI within the technology, eliciting reasoning around how the technology realises the HDI tenets effectively (or not), which of these decisions they would subscribe to and what they think should be changed. Each group would be given a starter kit which would include details of the technology, aspects of HDI included within it and how the makers intended for these decisions to alleviate socio-technical challenges, that would give the detail required to produce the critique.

5.5   Activity 4 : Rebuild : Break and Make [30 minutes]

Prepared by the information from the showcase and their discussions in the previous sessions, the respondents would now, in their groups, have the chance to rebuild data-driven technology ideas/ prototypes that respond to the socio-technical challenges by applying the principles of HDI. Here, each group would be given a popular data-driven technology, backed by a story of data abuse or breach (a mix of fiction and popular media stories). The groups will be given the opportunity to unpack the challenges presented in these technology-driven scenarios and see how they could intervene here with HDI. They would be asked to 'break' these current technological norms and rebuild them by 'making' new HDI-led responses that make these technologies less susceptible to the challenges.

5.6   Break [15 minutes]

5.7   Activity 4 (Contd.) : Rebuild : Break and Make [15 minutes]

5.8   Activity 5 : Feedback and Discussion [45 minutes]

This session would have every group present the following to the entire room :

1. Their technology-story
1. The HDI-inspired design revisions they made here
2. Challenges they faced during inclusion of these revisions
3. Highlighting any interplay of what needs to be done versus what can be done
4. An account of how this situation is better because they included HDI (Or whether it would have been better without HDI?)

Here, the wider group would be given the opportunity to raise comments, questions and queries, if required.

5.9   Activity 6 : Wrap up Discussion [30 minutes]

In this final session, the entire group would be brought back in order to reconsider the core research questions that the workshop sought to address : Are the current HDI concepts still the right ones as prevalence of complex systems have emerged? How do new human-AI interactions challenge the tenets, given the complexity and "black-boxing" of such systems? How could/ should HDI's tenets be refined to promote impact?

The aim of this discussion will be to seek more granular and nuanced framings of, and responses to, these questions we pose. These questions will form the basis for the post workshop plans (including a report and proposal for an ACM TOCHI Special Issue), and become the keystones in a new research agenda for HDI.

6 DISTANCE ENGAGEMENT

As in previous recent workshops we conducted at other ACM conferences, we would be leveraging an online platform to allow for presentation, discussion, feedback and other group-activities. We believe, for the purposes and processes planned for this workshop Gather[1] would be an appropriate choice. Gather is an alternative approach to video chat which uses characters in a 2D environment, along with a proximity heuristic, to allow for more natural video-based conversations. In particular, it allows breakout groups to form and reshape organically during workshop activities, and in the gaps between scheduled activities participants can easily talk to each other privately or in small groups while remaining visible and 'part' of the broader workshop space (replicating how 'coffee break chats' tend to work offline). Gather also enables the easy choice between emojis and avatars or video, instead of a mandatory camera-on protocol thereby allowing all participants inclusion in the sessions and social activities without having to compromise on their individual priorities around video use. Gather will act as the virtual 'space' that the event takes place within, with additional tools such as Miro (which seamlessly integrate with Gather) being used to facilitate synchronous creative collaboration on workshop tasks.

We would have 'assigned organisers' dedicated to managing the technicalities and flow of the online groups, helping participants with the discussions, activities and presentation of their feedback to ensure inclusion of

---

[1] https://gather.town/

their participation and contribution to the fullest. Participant accessibility requirements will be recorded right from the application stages and suitable adjustment measures managed on an individual basis before and during the workshop. For example, if captioning is required a suitable video 'proxy' will be set up to work alongside Gather. While the option of streaming the workshop itself for wider viewing is possible, given the participatory nature of the latter sessions, we believe it best to limit the number of participants to a maximum of 25 for optimum inclusion and engagement (see below for specific post workshop engagement plans).

7 POST WORKSHOP PLANS

The aim of the workshop is to initiate dialogue around the use of design to bridge the gap between the conceptual underpinnings of HDI and its tangible realisation within everyday technologies. Given the fledgeling nature of this area, there is the possibility to explore this avenue using different methods, within varying contexts and from diverse perspectives. Hence, we are in conversation with the ACM TOCHI team for the approval of a Special Issue where the conversations and collaborations initiated at the workshop could be further worked upon and published to contribute to this emergent field. While the organisers' networks and contacts would support this initiative, the execution and success of the workshop and engagement from the international CHI community would play an integral role in this process, which would be a significant milestone in broadening and furthering the scope and reach of HDI as it stands today.

The outcomes of the workshop would also be published as a report which would be made publicly available via the workshop website, and shared with our industrial project partners (e.g. BBC, Microsoft Research). This would contribute to the next generation of research around HDI within academia [8], contribute to research teams within industry (eg., BBC R&D's HDI initiative [5] and New Forms of Value initiative [4]) and also be part of the knowledge base for HorizonDER who are consistent contributers to policy making, particularly around the socio-technical implications of personal data leverage, within the U.K. [15].

8 CALL FOR PARTICIPATION

We welcome you to our workshop exploring responses to socio-technical challenges of personal data and AI through the application of Human Data Interaction. The workshop will: orient you to the emergent and multi-disciplinary field of HDI; present data driven experiences that have inculcated the tenets of HDI (legibility, negotiability, agency); welcome your critique of these projects; encourage you to imagine how HDI's tenets may best be realised.

For participation, you can either upload a two-page position paper (ACM Master Article Submission Template) on the workshop website : this could either describe an example of employing HDI, explain why HDI was used, what it helped/ failed to achieve and how other research may benefit from using it. Alternatively, you could offer a critique on HDI's tenets - using examples, theory or rhetoric - their strengths/ limitations and why these are the right/ wrong concepts. You could also fill out an online form on the workshop website, a more informal method of expressing your views on the same topics. A diverse group of complementary participants will be selected based on their prior experience of HDI (or related areas) and the quality/ relevance of their submission to the workshop theme.

Upon acceptance, it is mandatory that at least one author of each accepted position paper must attend the workshop and that all participants must register for both the workshop and for at least one day of the conference. More details are available at [Website].

9 REFERENCES

[1] Charles Arthur. 2011. Sony suffers second data breach with theft of 25m more user details. *The Guardian*. Retrieved July 6, 2019 from http://www.guardian.co.uk/technology/blog/2011/may/03/sony-data-breach-online- entertainment.

[2] BBC. 2013. Sony fined over "preventable" PlayStation data hack. Retrieved from http://www.bbc.co.uk/news/technology-21160818

[3] BBC. 2019. Why the BBC does not want to store your data - BBC News. *BBC*. Retrieved August 15, 2019 from https://www.bbc.co.uk/news/technology-48825413

[4] BBC. 2019. New Forms of Value: A BBC for the data economy. Retrieved October 12, 2020 from https://www.bbc.co.uk/rd/projects/new-forms-value-bbc-data-economy

[5] BBC R&D. 2017. Human Data Interaction - BBC R&D. Retrieved August 6, 2020 from https://www.bbc.co.uk/rd/projects/human-data-interaction

[6] Carole Cadwalladr and Emma Graham-Harrison. 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. Retrieved August 28, 2018 from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[7] UK Parliament Committees. 2020. Formal meeting (oral evidence session): Online harms and the ethics of data. Retrieved October 13, 2020 from https://committees.parliament.uk/event/2266/formal-meeting-oral-evidence-session/

[8] HDI. 2020. Human Data Interaction – A UK EPSRC Network Plus. Retrieved October 14, 2020 from https://hdi-network.org/

[9] Joseph Lindley. 2019. Gogglebox of the Future 2019 Part 2 - Reflections - YouTube. Retrieved October 14, 2020 from https://www.youtube.com/watch?v=w3zULRNLEH8

[10] Joseph Lindley, Ali Akmal, Fraziska Pilling, and Paul Coulton. 2020. *Researching AI Legibility through Design*. Retrieved August 6, 2020 from https://eprints.lancs.ac.uk/id/eprint/140343/1/Researching_AI_Legibillity_through_Design_pre_print.pdf

[11] Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek Mcauley, and Jon Crowcroft. 2014. Human-Data Interaction : The Human Face of the Data-Driven Society.

[12] Neelima Sailaja. 2020. Understanding the Challenges of Using Personal Data Use in Media Experiences (Currently In Minor Corrections).

[13] Neelima Sailaja, Andy Crabtree, James Colley, et al. 2019. The Living Room of the Future. *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, ACM, 95–107. Retrieved November 19, 2019 from https://dl.acm.org/citation.cfm?id=3323360

[14] Neelima Sailaja, Andy Crabtree, Derek Mcauley, and Phil Stenton. 2018. Explicating the Challenges of Providing Novel Media Experiences Driven by User Personal Data. *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, ACM.

[15] Neelima Sailaja and Derek McAuley. 2020. *Written evidence submitted by Neelima Sailaja and Prof Derek McAuley Response to the Call for evidence The future of Public Service Broadcasting Specific Response to "Looking ahead."* Retrieved October 12, 2020 from https://committees.parliament.uk/writtenevidence/7071/pdf/

[16] Neelima Sailaja, Derek Mcauley, Adrian Gradinar, Paul Coulton, and Rhianne Jones. 2020. Designing for Data Legibility, Negotiability and Agency in Personal Data Driven Media Experiences. *Proceedings of the 2020 ACM International Conference on Interactive Media Experiences*, ACM. Retrieved from https://www.design4data.info