



HAL
open science

The Human(s) in the Loop-Bringing AI and HCI Together

Tom Gross, Kori Inkpen, Brian Y. Lim, Michael Veale

► **To cite this version:**

Tom Gross, Kori Inkpen, Brian Y. Lim, Michael Veale. The Human(s) in the Loop-Bringing AI and HCI Together. 17th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2019, Paphos, Cyprus. pp.731-734, 10.1007/978-3-030-29390-1_71 . hal-02878630

HAL Id: hal-02878630

<https://inria.hal.science/hal-02878630v1>

Submitted on 23 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

The Human(s) in the Loop— Bringing AI and HCI Together

Tom Gross¹, Kori Inkpen², Brian Y. Lim³, Michael Veale⁴

¹ University of Bamberg, Bamberg, Germany, email(at)tomgross.net

² Microsoft Research, Redmond, WA, USA, kori(at)microsoft.com

³ National University of Singapore, Singapore, brianlim(at)compnus.edu.sg

⁴ University College London, London, UK, m.veale@ucl.ac.uk

Abstract. This document presents the motivation, objectives and target audience, theme, submissions, structure as well as the expected outcome.

Keywords: HCI, Artificial Intelligence, AI, Human in the Loop.

1 Motivation

As more and more artificial intelligent systems become incorporated into our everyday lives, it is critical that we understand the ways in which people will interact with these systems. Although some AI systems will be fully automated, a large number will be incorporated into a larger social ecosystem where people will be interacting with these systems. In some cases, advances in machine learning are enabling systems to make inferences on data that are more precise than human experts, however, there is also a growing body of literature that shows that these systems have inherent bias and can have a negative impact on human decision making [3]. It is imperative that researchers understand the smart interplay of AI systems and human experts such that the combination of the two can leverage the inherent strength and weaknesses of each to lead to optimal results. In this workshop, we seek to bring together researchers from both Artificial Intelligence and Human-Computer Interaction communities to discuss concepts, systems, designs, and empirical studies focusing on the communication and cooperation between individual users and teams of users with AI systems.

2 Objectives & Target Audience

The goal of this workshop is to bring together researchers from diverse communities such as Human-Computer Interaction, Machine Learning, Computer-Supported Cooperative Work, Interaction Design, Group Decision Support Systems, Visualisation, Philosophy and Ethics. It will build on the insights gained from a larger workshop we are organising at CHI 2019 (<http://aka.ms/whereisthehuman>), but focus in more specifically in issues related to Human+AI interaction.

The structure will focus on creating research partnerships and identifying collaborative projects. Although we will spend time discussing key trends, challenges, and opportunities, the overall goal is to have a focused workshop that will initiate projects that will extend beyond the workshop itself. The intimate nature of workshops at INTERACT is an ideal venue for this type of workshop.

3 Theme

This workshop will focus on three sub-themes related to Human + AI Collaboration:

1. **Integrating Artificial and Human Intelligence:** AI systems and humans both have unique abilities and are typically better at certain complementary tasks than others. For instance, while AI systems can summarize voluminous data to identify latent patterns, humans can extract meaningful, relatable, and theoretically grounded insights from such patterns. What kind of research designs are most amenable to and would benefit the most from combining artificial and human intelligence? What challenges might surface in attempting to do so? How do issues of trust and accountability impact results [5, 7]?
2. **Collaborative Decision Making:** How can we harness the best of humans and algorithms to make better decisions than either alone? How do we ensure that when there is a human-in-the-loop—such as in complex or life-changing decision-making—they remain critical and meaningful, while creating and maintaining an enjoyable user experience? Where is the line between decision support anticipating the needs of the user and it removing the user’s ability to bring in novel, qualitative critical knowledge to enable the system’s goals?
3. **Explainable and Explorable AI:** What does the human need to effectively utilize AI insights? How can users explore AI systems’ results and logic to identify failure modes that might not be easy to spot? Examples might be undesirable impacts on latent groups not corresponding to categories in the dataset [6], difficult-to-spot changes (‘concept drift’), or feedback loops in the socio-technical phenomena the AI system is modelling over time [2].

4 Submissions

Selection of participants and presentations will be based on refereed submissions. We invite authors to submit 4-page papers reporting their contributions in the field of the workshop or 2-page position statements motivating their interest in specific workshop topics. Papers should be formatted according to the INTERACT 2019 (Springer LNCS Series) format. An expert panel of 3-4 researchers will be recruited to review the submissions and participate in the conference.

Authors of accepted submissions as well as invited researchers will give short presentations. We will interweave presentation sessions with longer periods of discussions. Presentations will be grouped by key topics to foster spontaneous discussions.

If participants are interested, archival publication opportunities will be discussed, in addition to follow-on workshops.

5 Proposed Workshop Structure

0900 - Welcome and Introduction
0915 – Lightning Talks by workshop participants
1015 - Mid-morning break
1030 – Full-group brainstorming of possibly project areas
1200 - Lunch break
1300 - Breakout Groups
1430 - Mid-afternoon break
1500 - Report back from Breakout Groups
1600 - Brainstorm next steps
1700 - Workshop concludes

6 Expected Outcomes

Three key outcomes are expected from this workshop. First, community building and networking among key researchers in the area of Human+AI collaboration, with the potential to lead to future collaborations on projects or larger grant proposals. Second, an outline of important research directions for this emerging area. Third, one or more research projects that will continue beyond the workshop, the results of which will be published in premiere research venues.

7 Organisers

All organisers have successfully organised workshops at various scientific events of HCI and CSCW individually and together.

Dr. Tom Gross is full professor and chair of Human-Computer Interaction at the University of Bamberg, Germany. His research interests are particularly in the fields of Computer-Supported Cooperative Work, Human-Computer Interaction, and Ubiquitous Computing. He has participated in and coordinated activities in various national and international research projects and is a member of the IFIP Technical Committee on ‘Human Computer Interaction’ (TC.13). He has been conference co-chair and organiser of many international conferences. Further information can be found at: <http://www.tomgross.net>.

Dr. Kori Inkpen is a Principal Researcher at Microsoft, where she is a member of the Microsoft Research AI team. Dr. Inkpen’s research interests are focused on Human+AI Collaboration to enhance decision making, particularly in high-impact social contexts which inevitably delves into issues of Bias and Fairness. Kori has been a core member of the CHI community for over 20 years. Prior to joining Microsoft she was a Professor of Computer Science at Dalhousie University and Simon Fraser Uni-

versity. Further information can be found at: <http://research.microsoft.com/en-us/people/kori>.

Dr. Brian Y. Lim is an assistant professor in the Department of Computer Science at the National University of Singapore. He is leading the NUS Ubicomp Lab, where he and his team design, develop, and evaluate needs-driven infocomm technologies to address new societal challenges, such as urban systems, sustainability and energy management, healthcare and well-being. He has conducted research in intelligent systems across multiple modalities (IoT sensors, mobile interfaces, web and dashboards) and multiple scales (smartphones, smart homes, and smart cities). This allows me to develop impactful technological solutions for multiple domains, and to translate these innovations from the lab to society. Further information can be found at: <http://www.brianlim.net/>.

Michael Veale is a doctoral researcher in responsible public sector machine learning at the Dept. of Science, Technology, Engineering & Public Policy at University College London. His work spans HCI, law and policy, looking at how societal and legal concerns around machine learning are understood and coped with on the ground. His work on the governance of data-driven technologies has been debated in Parliament; cited by regulators and utilised by a wide range of international civil society groups and think-tanks. He acts as a consultant to a range of national and international governments working to ensure that public values are reflected in public sector technologies. Michael sits on the Advisory Council of the Open Rights Group and is a technical advisor on machine learning to the Red Cross Red Crescent Climate Centre. Further information can be found at <https://michaev.liv.ac.uk/>

References

1. Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Shu, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R. and Horvitz, E. Guidelines for Human-AI Interaction. In Proceedings of the Conference on Human Factors in Computing Systems - CHI 2019.
2. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. A survey on concept drift adaptation. *ACM Computing Surveys* 1, 1. <https://doi.org/10.1145/2523813>
3. Green, B., and Chen, Y. Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments. In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency - FAT* 2019.
4. Gross, T. Supporting Informed Negotiation Processes in Group Recommender Systems. *i-com - Journal of Interactive Media* 14, 1 (Jan. 2015). pp. 53-61.
5. Saxena, N.A., Huang, K., DeFilippis, E., Radanovic, G., Parkes, D.C., Liu, Y. How Do Fairness Definitions Fare? Examining Public Attitudes Towards Algorithmic Definitions of Fairness. In Proceedings of Association for the Advancement of Artificial Intelligence – AAAI 2019.
6. Veale M. and Binns, R. Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society* 4, 2. <https://doi.org/10/gdcfnz>
7. Yin, M., Wortman Vaughan, J., Wallach, H., Understanding the Effect of Accuracy on Trust in Machine Learning Models. In Proceedings of the Conference on Human Factors in Computing Systems - CHI 2019.