

Ontika et. al., (2022): Exploring Human-Centered AI in Healthcare: Diagnosis, Explainability, and Trust. In: Proceedings of the 20th European Conference on Computer Supported Cooperative Work: The International Venue on Practice-centered Computing on the Design of Cooperation Technologies - Workshops, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.48340/ecscw2022_ws06

Copyright 2022 held by Authors, DOI: 10.18420/ecscw2022_ws06

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

Exploring Human-Centered AI in Healthcare: Diagnosis, Explainability, and Trust

Nazmun Nisat Ontika¹, Hussain Abid Syed¹,
Sheree May Saßmannshausen¹, Richard HR Harper², Yunan Chen³,
Sun Young Park⁴, Miria Grisot⁵, Astrid Chow⁶, Nils Blaumer⁷,
Aparecido Fabiano Pinatti de Carvalho¹, Volkmar Pipek¹

¹ Institute for Information Systems, University of Siegen, Germany

² The Institute for Social Futures, Lancaster University, UK

³ Department of Informatics, University of California, USA

⁴ School of Information, University of Michigan, USA

⁵ Department of Informatics, University of Oslo, Norway

⁶ Eleanor Health, USA

⁷ Gemedico, Germany

{nazmun.ontika, hussain.syed, sheree.sassmannshausen, fabiano.pinatti, volkmar.pipek}@uni-siegen.de, r.harper@lancaster.ac.uk, yunanc@ics.uci.edu, sunypark@umich.edu, miriag@ifi.uio.no, astrid.chow@eleanorhealth.com, blaumer@gemedico.com

Abstract. AI has become an increasingly active area of research over the past few years in healthcare. Nevertheless, not all research advancements are applicable in the field as there are only a few AI solutions that are actually deployed in medical infrastructures or actively used by medical practitioners. This can be due to various reasons as the lack of a human-centered approach for the or non-incorporation of humans in the loop. In this workshop, we aim to address the questions relevant to human-centered AI solutions associated with healthcare by exploring different human-centered approaches for

designing AI systems and using image-based datasets for medical diagnosis. We aim to bring together researchers and practitioners in AI, human-computer interaction, healthcare, etc., and expedite the discussions about making usable systems that will be more comprehensible and dependable. Findings from our workshop may serve as 'terminus a quo' to significantly improve AI solutions for medical diagnosis.

Introduction

Artificial intelligence (AI) technology is highly demanded these days in healthcare, with the potential to empower healthcare professionals in their decisions, by providing them relevant information at a glance when it matters the most. The grim fact is that there are insufficient specialist physicians to fulfill the increasing demand for healthcare (IHS Markit Ltd., 2021). However, by incorporating AI into healthcare, we can possibly assist physicians to become more productive and efficient, helping to make informed choices easily and quickly. AI technology has great potential in decision-making because it is able to process a large amount of data in a short period of time, thus quickly providing medical professionals with a pre-diagnosis which builds the capacity to enhance the final judgements (Alsagheer et al., 2021).

Despite the great recent developments in AI across many sectors, especially healthcare, just a fraction of AI systems has effectively transitioned from the laboratory to medical practice. The absence of a human-centered approach while developing the systems, the complexity and unreliability of the final applications, the failure to include people in the development loop, and the lack of explainability for practitioners are frequently depicted as the key obstacles for greater AI adoption, but with reasonable reasons (Abdul et al., 2018). Different components within an organizational infrastructure are integrated through standardized interfaces enabling the work practitioners to channel merits like reflexivity, longevity, resilience, and heterogeneity (Hanseth & Lundberg, 2001; Pipek & Wulf, 2009; Syed et al., 2021). Medical organizations and practitioners, we argue, would have a difficult time dealing with AI if it does not integrate effortlessly into their present infrastructure, or even worse if it adds more complications. Moreover, any new technology can be difficult to develop, and even more difficult to gain trust when having a strong infrastructure as healthcare, where physicians must take immediate choices with foreseeably many further implications. Numerous endeavors to design usable systems for physicians fail due to insufficient task analysis, in which essential needs are either not discovered or their importance is undervalued (Preim & Hagen, 2011). Despite AI having demonstrated great promise in healthcare and medicine, so much of this development has not been implemented because machine learning models are trained on quasi data and

assessed only in controlled experiments that are quite different from real-world implementation circumstances (Okolo, 2022).

Human-AI systems operating jointly, rather than alone, have a great potential for high effectiveness (Ahuja, 2019; James Wilson & Daugherty, 2018). Since the relationship between a person's faith in automation and the capability of automation differs (J. D. Lee & And Moray, 1994; Muir, 1987), it leads to overtrust and distrust (J. D. Lee & See, 2004). Several pieces of research have revealed how medical practitioners overlooked suggestions identified by AI (de Boo et al., 2009; Nishikawa et al., 2006) and, also missed anomalies that AI failed to identify (Jorritsma et al., 2015) because of distrust and overtrust of AI, respectively. Adoption requires trust, which is difficult to earn (Pipek & Wulf, 2009). On the one hand, AI excelled physicians in identifying breast cancer quicker and better (Killock, 2020), but on the other hand, certain AI systems fared badly and might unwittingly cause more damage than good if used to influence treatment decisions (Wynants et al., 2020). Hence, for the diagnosis and treatment support, practitioners should remain in a position to make the final decisions. AI should aid them in pre-analysis rather than taking over the ability to make decisions.

Humans must be placed at the heart of AI development lifespans (Harper, 2008; Inkpen et al., 2019). Real users should be at the forefront and must be engaged with the system from the very beginning. We must go beyond the technology to comprehend the entire context of use. The fundamental pain points will be illdefined unless we grasp the true needs of the users through field investigations (Wulf et al., 2018). This would allow the creation of dynamic learning systems by keeping humans-in-the-loop (Syed et al., 2020). Furthermore, users must be aware of what an AI system can and cannot accomplish, as well as on what data it has been trained and for what it has been optimized. AI systems have gotten considerably more reliable and robust because of the emergence of deep learning, but they have also become somewhat difficult to comprehend (Ribeiro et al., 2016). The AI system's black-box aspect can be a barrier to credibility, which is why we must overcome the hurdle of interpretability by making AI's discoveries as apparent as feasible and necessary (Goldstein et al., 2015; Molnar, 2022; Wachter et al., 2017). We need to observe the practitioners in using the system to understand how mental models evolve with every success and failure, but also how the AI systems influence their decisions (Green & Chen, 2019).

Value Sensitive Design (VSD) (Friedman et al., 2006); is an approach that can be helpful to address the issues above. According to its premises, designers should design all technologies in a principled and comprehensive manner, accounting for human values throughout the design process (Friedman & Hendry, 2019). Human values such as fairness and responsibility should be included early in the design process to help designers apply their design abilities wisely. VSD may also assist in identifying aspects in technology that promote, impede, or prohibit certain values once the effects and significance of the selected values are recognized. VSD

encourages us to think about human values as a design requirement in the same way that we think about efficiency, effectiveness, usability, accessibility, and dependability (Davis & Nathan, 2015). It has been argued that VSD will keep expanding and shape the future way of thinking while designing solutions (Friedman et al., 2017; Umbrello & de Bellis, 2018).

Performance and explainability are currently on the table as a trade-off. Models with the highest performance (e.g., deep learning) are generally the least explainable, whereas models with the worst performance (e.g., linear regression, decision trees) are often the most explainable (Kelly et al., 2019). The true objective of Explainable AI (XAI) ought to be to guarantee that end-users can perceive the results, thereby assisting them to enhance their decision-making performance (Gunning et al., 2019). Researchers in charge of building explanatory user interfaces should be involved in the development of Human-centered AI (HAI) technologies (Shneiderman, 2020). Moreover, end-users need to be engaged in the development of such explanation interfaces. The interface should supply descriptions for any algorithmic decisions, but it should also supply various layers of rationalizations, allowing the end-user to question the AI decision-making operation, possibly down to the stage of any datasets used in the machine learning development in exploring the complete data origin and its boundaries (Xu, 2019). Engaging people in the design and putting them in the limelight enhances the likelihood that the resulting systems will be ethical, adaptable, useful, and deployed, and that adverse unexpected effects of AI systems are avoided (Bond et al., 2019).

In terms of adoption, it is sensible to think that XAI techniques are likely to speed up the adoption of AI solutions in medical environments while also fostering crucial transparency and trust with potential users, since any mistakes might not only affect the patients but also impede the use of such solutions (Adadi & Berrada, 2018). Here, visualization becomes an important aspect. Through visualizations, the decision of the AI system is made more understandable and transparent, which leads to a fair and responsible perception (M. K. Lee, 2018). This, in turn, can lead us forward in a discussion of whether visualization can be the first step of explainability because we need to find more techniques to represent medical knowledge more meaningfully.

Graphical user interface design is also a critical aspect to consider for any product development. Healthcare physicians, as users in general, want user interfaces that are simple to operate yet aesthetic, as well as intriguing and encouraging (Wang et al., 2021). In order to get to this point, qualifications that go beyond visualization techniques are necessary (Preim & Hagen, 2011). Hence, collaborations with researchers in the field of human-computer interaction (HCI) are strongly encouraged, with especial attention to issues from psychology, visual design, and user interface design, etc. These interdisciplinary collaborations can

potentially lead to more useful and usable advanced user-centered medical visualization mechanisms.

This workshop provides an incubator for the researchers and practitioners to create a joint consortium for research towards a wide range of practices and technologies. The workshop offers a great opportunity to instigate the discussion about shortening the gap in Medical AI and human-centered design. Emerging ideas will be further pursued in future publication plans. Our workshop targets contributions showing how different HCI approaches for XAI have been used in current and past research and field works and aim at reflecting on the lessons learned from them. Incorporating HAI into healthcare effectively is a significant venture with constraints that entail a multi-disciplinary approach combining specialists from HCI, AI, healthcare, psychology, and social sciences. This workshop will address important HAI concerns, enabling optimal human-machine integration by enhancing the trustworthiness between humans and technology. We will discuss ways to assure that AI applications focus on the end-user, put humans in the loop, and emphasize human values in a responsible manner. We will explore different prototyping and evaluation techniques; also, how we can integrate the context of use with real user needs and usage scenarios into task analysis methods; and how all these can help make new strategies to improve the overall user experience.

Workshop Goals and Topics

The goal of this workshop is for participants to explore various approaches of Human-Centered AI and to develop a strategy for future scientific investigations on healthcare solutions. We will use a cross approach, gathering different points of view together to discuss the numerous benefits and drawbacks of such innovations. We would also like to learn from other fields and approaches, that are developing and using AI with visualizations in similar contexts with a human-centered approach.

We hope to address the following themes and questions but are not limited to:

Workshop Themes:

- Human-Centered AI for medical visualizations
- Physician-in-the-loop for HAI
- Explainable AI in healthcare
- Trust and fairness issues of AI in healthcare
- Ethics in AI for healthcare
- Security and privacy in medical AI

Research Questions:

- What are the existing human-centered approaches for designing an AI-based medical diagnosis?
- How are end-users integrated into the development process?
- How is it possible to make AI decisions comprehensible and transparent to the end-user?
- What are other examples/ use cases, in which image-based detection/ diagnosis is done?
- What is the role of visualization in XAI?

Participation

Our two-half-day workshop will be held in person, providing that the current pandemic situation allows. However, we will also provide alternatives for participants who cannot attend it in person - e.g., through a Zoom link. The infrastructure for the in-person workshop will be provided by the conference and secured by the workshop organizers. Participants attending online will be responsible for arranging the necessary equipment, namely a computer, video camera (external or integrated with laptop), microphone, paper, pen, etc. to attend the workshop. However, the organizers will support the participants for any technical troubleshooting (e.g., handling presentations on Zoom, doing activities on Miro, etc.) during the sessions. Considering attendees will be engaging on microphones and cameras within an academic community during the interactive workshop, it will necessitate a private place, free of unwanted distractions and disruptions.

We will invite researchers and practitioners from academia and industry pursuing research about HCI, AI, HAI, XAI, and Healthcare Informatics. A call for contributions will be sent. The organizers will also directly contact different communities and relevant social media outlets. Through distribution lists, social media, and personal contacts, people with industry expertise and interest in adjacent sectors will be approached. All information on the workshop, including the workshop themes, submission process, and important deadlines are available on our workshop website available at <https://ecscw2022-hcai.yolasite.com>.

A maximum of 10 position papers and 20 participants, excluding the organizers, will be admitted for the workshop to provide a more structured discussion and increase the likelihood of achieving useful outcomes. The workshop will require a minimum of one author from each accepted paper to register and attend. To participate actively in discussions, all participants are encouraged to read the workshop contributions, which will be accessible prior to the workshop.

Submission and Selection

Workshop participants will be asked to submit a position paper following the ECSCW template 2-4 pages including bibliography, short use-cases: presenting materials, ideas, AI technology or artifacts they would like to discuss in the workshop. Submissions should include a brief outline of the main ideas and arguments for the contribution. Participants can also submit case-studies or reports on the recent experiments in their research context, prototypes, demos, or other research formats, they would like to demonstrate or discuss during the workshop. Participants who wish to contribute to the discussion without submitting material are not required to submit a position paper. The submission and the review process will be managed over e-mail. Workshop participants must submit a position paper by the deadline to hai.health.ecscw2022@gmail.com. The submissions should not be anonymized and will be reviewed by the workshop organizers and selected based on their quality, consistency with the workshop theme, and potential to generate fruitful discussions during the workshop.

Important Dates

- April 22nd, 2022: Submission of position papers
- May 6th, 2022: Notification of acceptance
- June 3rd, 2022: Camera-ready
- June 27th - 28th 2022: Workshop Days

Workshop Structure

The hybrid workshop will be held on two consecutive days, on June 27th and 28th with three hours each day including short tea break in between, within conference preferred timeslots, 14:00-17:00 UTC+1. The tentative event structure of our two-day interactive workshop is (roughly) as follows:

Workshop initiation: The organizers will initiate the proposal, laying out the workshop's objectives, goals, and anticipated benefits in detail. The participants will briefly introduce themselves during this session.

Interactive case study analysis: Participants will showcase the material they bring to the discussion. All the participants will be asked to engage through questions and answers. A minimum of 25 minutes will be dedicated to each case study. This time can be slightly longer, if less than 10 position papers are accepted. This activity is designed to engage the gathering with personal observations and to stimulate conversation on subjects that will be discussed in future sessions. Naturally, this will not allow for in-depth study of the cases, and this is also not the

intention. Instead, we intend to increase group topic motivation while identifying substantial discussion themes.

Interactive brainstorming session: We will next proceed by selecting problems for further discussion as a group. We will break off into smaller groups for discussion. A smaller number of organizers will moderate each group, which will be given a theme to discuss upon. The topics will be examined in further depth, this time using the example set of case studies to investigate the many issues that arise.

Plenary session: Following the group work, we will gather as a group and report briefly on the various conversations and conclusions.

Wrap-up: The organizers will provide closing comments and highlight the workshop's key lessons.

The organizers will also address the idea of teaming with the participants on a collaborative publication to make the findings available to the CSCW, HCI and AI research community.

Organizers

Nazmun Nisat Ontika, M.Sc., is a Research Assistant at the chair of CSCW and social media at the University of Siegen. Her research interests of late include Human-computer Interaction, User-Centered Technology Design, Child Computer Interaction, Virtual and Augmented Reality for better Accessibility and Usability. Her current research includes Human-Centered Artificial Intelligence in Radiology.

Hussain Abid Syed, M.Sc., is a Ph.D. scholar at the chair of CSCW and social media at the University of Siegen. His research interests are in crisis informatics, infrastructures, explainable AI, and human centered AI. His current research includes exploring the phenomena of organizational resilience and infrastructuring in small and medium enterprises and developing lightweight socio-technical solutions using technologies like service-oriented architecture, rest APIs and data science pipeline.

Sheree May Saßmannshausen, M.Sc., is a Research Assistant at the chair of CSCW and social media at the University of Siegen. Her research interests are in the field of Human-Computer-Interaction and Human-Centered-AI in the context of healthcare. Her current research includes User Experience Design for technologies like Augmented Reality or Artificial Intelligence.

Richard HR Harper, Ph.D., is a Professor of Computer Science and Director of the Institute for Social Futures at Lancaster University. He is a Fellow of the IET, Fellow of the SIG-CHI Academy of the ACM, Fellow of the Royal Society of Arts,

and Visiting Professor in the College of Science at the University of Swansea, Wales. His research is primarily in Human Computer Interaction, though it also includes social and philosophical perspectives. His research on trust in HCI has ranged from explorations of file abstractions, the role of trust in the self, and how trust is a taken for granted feature of interaction. He has written 13 books, including ‘Trust, Computing and Society’ (Ed. CUP, 2015) the IEEE award winning ‘Myth of the Paperless Office’ (MIT: 2003)); and ‘Choice’ (Polity: 2016). He holds 26 patents, including ones for new cloud-based interaction devices (such as the ‘Cloud Mouse’), new secure data stores and lightweight mobile phone data exchange protocols. Prior to joining Lancaster, he was Principal Researcher at Microsoft Research.

Yunan Chen, Ph.D., is an Associate Professor of Informatics in the Donald Bren School of Information and Computer Sciences at the University of California, Irvine. Her research interests lie at the intersection of human-computer interaction (HCI), computer-supported cooperative work (CSCW), and health informatics. She is interested in data-driven technologies and human-centered AI for consumer health. Currently, she serves as Director of the Undergraduate Minor Program in Health Informatics, Vice Chair of Undergraduate Affairs at Department of Informatics, and Co-Director for the Health and Information Lab.

Sun Young Park, Ph.D., is an Associate Professor at the University of Michigan in the Stamps School of Art and Design and the School of Information. Her research lies at the intersection of Health Informatics, Human Computer Interaction (HCI), Computer Supported Cooperative Work (CSCW), Participatory Design, and Design Research. Her research uses design ethnography to study patient engagement, patient-provider collaboration, patient-centered health technology, and technology adaptation. Her work has been awarded by the National Science Foundation (NSF) and the Agency for Healthcare Research and Quality (AHRQ).

Miria Grisot, Ph.D., is an Associate Professor at the Department of Informatics, University of Oslo. Her main research interests are in the areas of information systems innovation, complexity and socio-technical systems, and organizational change, specifically in healthcare. She is engaged in research on AI in context. She is affiliated with the AI4users project addressing the “black box” problem contributing to the responsible use of AI for the digitalisation of public services. She is a member of the Association of Information Systems. She has worked mainly with an Information Infrastructure perspective, published in JAIS, CSCWJ, JSIS, SJIS.

Astrid Chow, M.S., MBA, is a Principal Product Designer and Strategist building the UX Design and User Research practice at Eleanor Health, a healthcare start-up

that focuses on substance use disorders, alcohol use disorders, and mental health. Astrid serves as a VP board member for the User Experience Professionals' Association (UXPA) Boston Chapter. Additionally, she is a frequent guest speaker and panellist on the topic of AI & Design Ethics at events such as the HFES Health Care Symposium, Connected Health conference, ACM CHI and CSCW conferences, and O'Reilly's AI Conference. She is teaching an elective course on Design Ethics in Practice for the University of Washington.

Nils Blaumer, M.A., is the Managing Director of Gemedico GmbH. His research interests are in workflow automation and digitalization of health processes. His current research includes the detection of prostate carcinomas in MRI images with the help of AI in order to considerably facilitate the daily routine of radiologists.

Aparecido Fabiano Pinatti de Carvalho, Ph.D., is an Associate Researcher at the Institute of Information Systems and New Media and Deputy Director of the Chair of Computer Supported Cooperative Work and Social Media, University of Siegen (Germany). His interests span human-computer interaction (HCI), computer supported cooperative work (CSCW), practice-centred computing, artificial intelligence (AI), software accessibility, cyber-physical systems, mobile and nomadic work, and informatics in education. The focus of his research is on technologically mediated human practices, more specifically on the understanding on how practices can help identifying the design space of new and innovative technologies, and how they can shape and be shaped by their usage. He has published several articles on topics related to these fields of research in prestigious international conferences.

Volkmar Pipek, Ph.D., is a Professor for Computer Supported Cooperative Work and Social Media with the Institute for Information Systems at the University of Siegen, Germany. He currently chairs to the board of trustees of the International Institute for Socio-Informatics (IISI). He has widely published books and articles in CSCW, with a specific interest in infrastructuring. He is also the co-leader of the project "INF - Infrastructural Concepts for Research in Cooperative Media" at the Collaborative Research Centre 1187: Media of Cooperation and the leader of the Project PAIRADS, which is a research project in the field of the integration of artificial intelligence in radiology at the University of Siegen.

Acknowledgments

The workshop organizers from the University of Siegen would like to acknowledge the financial support from the Bundesministerium für Bildung und Forschung - BMBF through the PAIRADS project (funding code: 16SV8651; <https://pairads.ai>).

References

- Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable and intelligible systems: An HCI research agenda. Conference on Human Factors in Computing Systems - Proceedings, 2018-April. <https://doi.org/10.1145/3173574.3174156>.
- Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- Ahuja, A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*, 2019(10). <https://doi.org/10.7717/peerj.7702>.
- Alsagheer, E. A., Rajab, H. A., & KM Elnajar, K. M. (2021). Medical Expert System to Diagnose the Most Common Psychiatric Diseases. The 7th International Conference on Engineering & MIS 2021, 1–6. <https://doi.org/10.1145/3492547.3492593>.
- Bond, R. R., Mulvenna, M., & Wang, H. (2019). Human centered artificial intelligence: Weaving UX into algorithmic decision making. *RoCHI 2019: International Conference on HumanComputer Interaction*.
- Davis, J., & Nathan, L. P. (2015). Value sensitive design: Applications, adaptations, and critiques. In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*. https://doi.org/10.1007/978-94-007-6970-0_3.
- de Boo, D. W., Prokop, M., Uffmann, M., van Ginneken, B., & Schaefer-Prokop, C. M. (2009). Computer-aided detection (CAD) of lung nodules and small tumours on chest radiographs. *European Journal of Radiology*, 72(2). <https://doi.org/10.1016/j.ejrad.2009.05.062>.
- Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press.
- Friedman, B., Hendry, D. G., & Borning, A. (2017). A survey of value sensitive design methods. In *Foundations and Trends in Human-Computer Interaction (Vol. 11, Issue 23)*. <https://doi.org/10.1561/1100000015>.
- Friedman, B., Kahn Jr., P. H., & Borning, A. (2006). *Value Sensitive Design and Information Systems (PREPRINT)*. *Human-Computer Interaction and Management Information Systems: Foundations*.
- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44–65. <https://doi.org/10.1080/10618600.2014.907095>.
- Green, B., & Chen, Y. (2019). Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. *FAT* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*. <https://doi.org/10.1145/3287560.3287563>.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI-Explainable artificial intelligence. *Science Robotics*, 4(37). <https://doi.org/10.1126/scirobotics.aay7120>.
- Hanseth, O., & Lundberg, N. (2001). Designing work oriented infrastructures. *Computer Supported Cooperative Work*, 10(3–4). <https://doi.org/10.1023/A:1012727708439>.
- Harper, R. (2008). Being human: Human-computer interaction in the year 2020. *Human-Computer Interaction*.
- IHS Markit Ltd., A. of A. M. C. (2021). *The Complexities of Physician Supply and Demand: Projections From 2019 to 2034*.

- Inkpen, K., Veale, M., Chancellor, S., de Choudhury, M., & Baumer, E. P. S. (2019, May 2). Where is the human? Bridging the gap between AI and HCI. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3290607.3299002>.
- James Wilson, H., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. In *Harvard Business Review* (Vol. 2018, Issue July-August).
- Jorritsma, W., Cnossen, F., & van Ooijen, P. M. A. (2015). Improving the radiologist-CAD interaction: Designing for appropriate trust. In *Clinical Radiology* (Vol. 70, Issue 2, pp. 115–122). W.B. Saunders Ltd. <https://doi.org/10.1016/j.crad.2014.09.017>.
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. In *BMC Medicine* (Vol. 17, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s12916-019-1426-2>.
- Killock, D. (2020). AI outperforms radiologists in mammographic screening. *Nat Rev Clin Oncol* 17, 134. <https://doi.org/https://doi.org/10.1038/s41571-020-0329-7>.
- Lee, J. D., & And Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. In *International Journal of Human-Computer Studies* (Vol. 40, Issue 1).
- Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1), 50–80.
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data and Society*, 5(1). <https://doi.org/10.1177/2053951718756684>.
- Molnar, C. (2022). *Interpretable Machine Learning A Guide for Making Black Box Models Explainable*.
- Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, 27(5–6). [https://doi.org/10.1016/S00207373\(87\)80013-5](https://doi.org/10.1016/S00207373(87)80013-5).
- Nishikawa, R. M., Edwards, A., Schmidt, R. A., Papaioannou, J., & Linver, M. N. (2006). Can radiologists recognize that a computer has identified cancers that they have overlooked? *Medical Imaging 2006: Image Perception, Observer Performance, and Technology Assessment*, 6146, 614601. <https://doi.org/10.1117/12.656351>.
- Okolo, C. T. (2022). Optimizing human-centered AI for healthcare in the Global South. *Patterns*, 3(2), 100421. <https://doi.org/10.1016/j.patter.2021.100421>.
- Pipek, V., & Wulf, V. (2009). Infrastructuring: Toward an integrated perspective on the design and use of information technology. *Journal of the Association for Information Systems*, 10(5). <https://doi.org/10.17705/1jais.00195>.
- Preim, B., & Hagen, H. (2011). HCI in Medical Visualization. In *Scientific Visualization: Interactions, Features*.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should i trust you?” Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*. <https://doi.org/10.1145/2939672.2939778>.
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction*, 36(6). <https://doi.org/10.1080/10447318.2020.1741118>.
- Syed, H. A., Schorch, M., Ankenbauer, S. A., Hassan, S., Meisner, K., Stein, M., Skudelny, S., Karasti, H., & Pipek, V. (2021). Infrastructuring for organizational resilience: Experiences and perspectives for business continuity. *Proceedings of 19th European Conference on Computer-Supported Cooperative Work*.

- Syed, H. A., Schorch, M., & Pipek, V. (2020). Disaster learning aid: A chatbot centric approach for improved organizational disaster resilience. Proceedings of the International ISCRAM Conference, 2020-May.
- Umbrello, S., & de Bellis, A. F. (2018). A value-sensitive design approach to intelligent agents. Artificial Intelligence Safety and Security (2018) CRC Press (. Ed) Roman Yampolskiy.
- Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3063289>.
- Wang, D., Wang, L., & Zhang, Z. (2021). Brilliant ai doctor in rural clinics: Challenges in aipowered clinical decision support system deployment. Conference on Human Factors in Computing Systems - Proceedings. <https://doi.org/10.1145/3411764.3445432>.
- Wulf, V., Pipek, V., Randall, D., Rohde, M., Schmidt, K., & Stevens, G. (2018). Socio-informatics: A practice-based perspective on the design and use of IT artifacts. In Socio-Informatics: A Practice-Based Perspective on the Design and Use of IT Artifacts. Oxford: Oxford University Press.
- Wynants, L., van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Bonten, M. M. J., Dahly, D. L., Damen, J. A., Debray, T. P. A., & others. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *Bmj*, 369.
- Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Interactions*, 26(4), 42–46. <https://doi.org/10.1145/3328485>.