

Panel 59. Disentangling AI and Health/Healthcare: imaginaries, Narratives, Values

Convenor:

Elisabetta Locatelli, Università Cattolica del Sacro Cuore

Keywords: AI, health, healthcare, imaginaries, values

Health and healthcare are among the fields with the most promising and innovative applications of AI technologies, for example for accelerating diagnoses; identifying rare diseases; interpreting diagnostic images; improving communication (i.e. if added in devices for disabled people or for dissemination purposes).

Besides the hype and the enthusiasm, a deep understanding is needed to unfold the hidden values and risks of AI for health and healthcare. Relevant issues include how fundamental concepts are conceived (such as health, healthcare, care, person, life, death, illness, efficiency); who should lead AI innovation; who is in charge to define boundaries and limits regarding the use of data (i.e. data collection and storage, avoiding biases, AI training). Beside the study of the technology per se, previous STS research demonstrated the importance of imaginaries and visions in shaping technology outcomes (Bijker, 1995; MacKenzie, Wajcman, 1999). Decoding such imaginaries would be helpful to better understand the issues above sketched out and direct the AI development for health and healthcare towards the common good.

Despite the number of investigations on AI imaginaries (well summed up by Sartori, Bocca, 2023; see also Bakiner, 2023), the association between AI and health/healthcare is still underinvestigated, as Hoff highlights in his analysis of how governments construe sociotechnical imaginaries of AI and healthcare services (Hoff, 2023; see also Tucker, 2023).

Beside governments, several other actors contribute to build this imaginary, such as innovators, companies, news media, institutional documents, popular discourses.

The panel proposes, thus, to put together different perspectives on the intersection between health/healthcare and AI, to investigate its technological, cultural, and sociological underpinnings, bias, and risks. Submissions of theoretical or empirical research regarding, but not limited to, the following topics are welcome:

- National and/or international imaginaries on AI and health/healthcare;
- Media (news media, social media...) imaginaries of AI and health/healthcare;
- Imaginaries of AI and health/healthcare from governments, public institutions, physicians, healthcare professionals, scientists, patients, patient's organizations, common citizens, etc.;
- AI for health/healthcare communication;
- STS analysis of design and implementation of AI for health/healthcare.

References

- Bakiner, O. (2023). Pluralistic sociotechnical imaginaries in Artificial Intelligence (AI) law: the case of the European Union's AI Act. *Law, Innovation and Technology*, 15(2), 558–582. <https://doi.org/10.1080/17579961.2023.2245675>.
- Bijker, W. E. (1995). *Of bicycles, bakelites, and bulbs: toward a theory of sociotechnical change*. MIT Press, Cambridge.
- Hoff, J. (2023). Unavoidable futures? How governments articulate sociotechnical imaginaries of AI and healthcare services. *Future*, 148: 1-13.
- MacKenzie, D., Wajcman, J. (1999). *The social shaping of technology*. Open University Press, Buckingham.
- Sartori, L., Bocca, G., (2023). Minding the gap(s): public perceptions of AI and sociotechnical imaginaries. *AI and Society*, 38: 443-458.
- Tucker, J. (2023). The future vision(s) of AI health in the Nordics: Comparing the national AI strategies. *Futures*, 149. <https://doi.org/10.1016/j.futures.2023.103154>.



11 JUNE 2025 09.00 - 11.00

ID 205 - Transformation of the Clinic. Data-driven pre-emption of disease and the politics of health

Niels Van Dijk, Vrije Universiteit Brussel

Katerina Sideri, Panteion University of Social and Political Sciences

Keywords: predictive medicine, artificial intelligence in healthcare, remote patient monitoring, healthcare innovation, actor-network theory

Public health policies are currently reimagining the function of healthcare systems in the light of a new explicit set of problematised challenges of an aging population with chronic disease, skyrocketing health expenses and costly treatments. Digital devices such as remote monitoring technologies and big data analytics are increasingly positioned as crucial nodes in enabling their transformation to address these issues. They are imagined by European regulators and funders to provide continuous real-time clinically relevant information about patient's everyday behaviour. In this line, these funders have announced calls to develop RMTs to enable data-driven pre-emption of episodes of chronic disease to avoid rehospitalisation of patients. This would allow more detailed preventive action, no longer merely based, in the words of Foucault, on 'the perpetual observation and measurement of the state of health of the population' through statistics, but that of the individual patient through concrete data-driven techniques of personalisation.

We will present the results of our case study of the work of a consortium of investigators of a European research project funded by the European Commission and pharmaceutical industry, which was seeking to develop such pre-emptive digital devices. We focus specially on their work of exploring ways in which these technologies could work in the clinic. The project here simulated and explored the transformations of the clinic and clinical action due to the introduction of these RMTs. Through exercises of 'interessement' of different actors in the clinical ecology, they solicited the views of clinicians, patients, healthcare managers and pharma companies regarding the 'value' their devices could bring. In this way they tried to make these digital devices the 'obligatory passage point' for innovation in healthcare. We argue that their proposals can be viewed as exercises in clinical world-making in action and thus as specific kinds of politics of health.

These exercises raise a series of questions regarding which actors are or are not included in these proposals and what different roles they are given, which actors reject these technologies and what kind of counterproposals they produce. We show how, throughout the stakeholder consultations, certain limits of these proposals become gradually apparent, and which other possible avenues are opened. In this way, these exercises also offer glimpses of other ways of 'world-making'. These counterproposals go hand in hand with diverse understandings of what healthcare is or should be about, what therapy is, and they would require a different organisation of the clinic and its link to the home, and of the patient-doctor relationship. As we show, the direction of the transformation of the clinic is still highly uncertain.

11 JUNE 2025 09.00 - 11.00

ID 288 - The risk of trust: promises and pitfalls of Artificial Intelligence in radiomics

Marta Gibin, Università di Bologna

Riccardo Pronzato, Università di Bologna

Antonio Maturò, Università di Bologna

Keywords: Artificial Intelligence, radiomics, trust, socio-technical imaginaries.

The introduction of artificial intelligence (AI) systems in healthcare has stimulated numerous debates regarding their potential benefits and risks, including their implications for professional roles, doctor-patient relationships, and trust. Different research areas, such as sociology of health (Lombi and Rossero, 2024), STS (Miele and Giardullo, 2024) and critical digital health studies (Lupton and Butler, 2024), have highlight-



ed the need to analyse these narratives in practice.

Within this framework, this paper aims to investigate narratives of risks, hopes and trust regarding the implementation of AI-based systems in healthcare and, more specifically, how people interpret and attribute meaning to the role of AI in breast cancer detection.

Narratives around technologies contribute to shaping the visions, fears and expectations people have around them, as well as personal and professional technological adoption (Bareis and Katzenbach, 2022; Sartori and Bocca, 2023). Specifically, a key dimension in the narratives surrounding AI is the "trust" towards it (Brown and Bahri, 2019; Hallowell et al., 2022), which can favour socio-technical and moral environments where new risk categories are created (Brown and Van Voorst, 2024). In this context, narratives have also been linked to "socio-technical imaginaries", i.e., an ambivalent term indicating both future and normative visions imbued in technological projects (Jasanoff and Kim, 2009), and end users' experiences (Bucher, 2017). Within this framework, we aim to show the contested and generative nature of AI narratives, and how "future individual expectations and trust" are key dimensions in "the study of technology put in context" (Sartori and Bocca, 2023: 446).

Given this scenario, this study provides insights into the visions and tensions surrounding the implementation of AI software programs to detect breast cancer. To do so, we conducted a content analysis of 701 online user comments responding to a New York Times article discussing AI use in breast cancer detection (Barker and Galardi, 2015). After several rounds of coding, we chose to apply Sztompka's trust framework (1995, 1999) to analyse AI narratives, which emerged as deeply interwoven with various forms of trust and distrust towards the medical profession, the U.S. healthcare system, the medical industry and the government. Specifically, findings reveal divergent visions of AI in healthcare, associated with four forms of trust identified by Sztompka: positional, segmental, organisational and technological trust. Findings show that positional trust in AI and physicians varies, with some viewing AI as complementary, others fearing dehumanisation and errors. Segmental trust in the profit-driven U.S. healthcare systems negatively influences attitudes toward AI, as does low organisational trust in medical institutions and governments. Technological trust both drives optimism and scepticism.

Our analysis underscores the need for a balanced approach to AI integration in healthcare, emphasizing collaboration between AI and physicians. Collaborative intelligence is proposed as a model for future practice. Such a framework would require physicians to develop technological competencies while maintaining their traditional skills in empathy and contextual interpretation.

11 JUNE 2025 09.00 - 11.00

ID 298 - Unraveling AI imaginaries in radiomics: Beyond exoticism, mentalism, and technologism

Jakub Mlynar, HES-50 Valais-Wallis University of Applied Sciences and Arts

Mélanie Champendal, HES-50 Valais-Wallis University of Applied Sciences and Arts

Lluís Borràs Ferrís, HES-50 Valais-Wallis University of Applied Sciences and Arts

Ricardo Ribeiro, HES-50 Valais-Wallis University of Applied Sciences and Arts

Adrien Depeursinge, HES-50 Valais-Wallis University of Applied Sciences and Arts

Keywords: AI, healthcare, radiomics, exoticism, mentalism, technologism

Sociotechnical imaginaries of 'AI' in healthcare are embodied in the design and practical implementation of AI-based technologies. Our paper is based on STS studies in radiomics (see, e.g., [1], for an overview), a new field of medical imaging analysis that involves extracting large-scale quantitative features using machine-learning (ML) algorithms. Specifically, we have been studying how radiomics is taught – formally and informally – among medical professionals and students at a university hospital in Switzerland, and how novice users interact with the radiomics platform QuantImage ([2]; [3]). The platform allows for the extraction of several types of features from PET/CT images, providing a simple environment that can be further adjusted for more refined analyses. It enables clinical researchers with no programming background to



develop and validate radiomics models using their own data, which can be easily exported from the hospital information system. QuantImage, initially designed as a tool for radiomics research, is currently being repurposed for education in this emerging field. Although the platform itself does not produce communicative actions, our findings show that autonomous agency is routinely ascribed by the participants to the machine, even though it functions as an object rather than an agent [4].

In this paper, building on our studies in radiomics and earlier work on AI imaginaries [5], we critically disentangle three aspects of sociotechnical imaginaries involved in teaching for and about AI in healthcare: exoticism, mentalism, and technologism. First, most of the existing insights and ways of working are obtained from settings in which AI has not yet been incorporated into the routine structures of everyday life, but is rather seen as a novel and unusual object (cf. [6]). The exoticist framing of AI in radiomics uncovers tension with the mundane work routines in which it is eventually embedded. Second, much of the social studies of AI reproduce (intentionally or not) the mentalist conception of AI, based on cognitivist notions of thinking, intelligence, or learning (cf. [7]). Such mentalism of AI in radiomics is connected to the imagery of an isolated single user taking part in individualised 'human–AI interactions'. Third, it is taken for granted that AI is a form of computational technology: AI and its manifestations – e.g., algorithms, neural networks or ML processes – are located inside the machine. Social imaginaries consist of technologist 'use cases/scenarios' that are produced to provide an optimal environment for the technology's operation. Against the backdrop of the three aspects, I will propose a reformulation of AI in healthcare that starts from the publicly observable, situated, embodied conduct in the world of daily life and professional activity, avoiding both the "academic and theoretical imperialism" [8] and the imperialism of computing [9]. More generally, I will argue that AI is not 'inside the machines' but emerges from the situated organisation of social events in which an agentic artefact's self-sufficiency is constituted and maintained.

References:

- Abler, D., Schaer, R., Oreiller, V., Verma, H., Reichenbach, J., Aidonopoulos, O., Evéquo, F., Jreige, M., Prior, J. O., & Depeursinge, A. (2023). QuantImage v2: A comprehensive and integrated physician-centered cloud platform for radiomics and machine learning research. *European Radiology Experimental*, 7(1), Article 16. <https://doi.org/10.1186/s41747-023-00326-z>
- Coulter, J. (2008). Twenty-five theses against cognitivism. *Theory, Culture & Society*, 25(2), 19–32. <https://doi.org/10.1177/0263276407086789>
- Guiot, J., Vaidyanathan, A., Deprez, M., Zerka, F., Danthine, D., & Leijenaar, R. T. H. (2022). A review in radiomics: Making personalized medicine a reality via routine imaging. *Medical Research Reviews*, 42(1), 426–440. <https://doi.org/10.1002/med.21846>
- Mlynář, J., de Rijk, L., Liesenfeld, A., Stommel, W., & Albert, S. (2024). AI in situated action: A scoping review of ethnomethodological and conversation analytic studies. *AI & Society*. <https://doi.org/10.1007/s00146-024-01919-x>
- Mlynář, J., & Arminen, I. (2023). Respecifying social change: The obsolescence of practices and the transience of technology. *Frontiers in Sociology*, 8, Article 1222734. <https://doi.org/10.3389/fsoc.2023.1222734>
- Mlynář, J., Bahrami, F., Ourednik, A., & Reichenbach, J. (2022). AI beyond Deus ex Machina: Reimagining intelligence in future cities. In *Proceedings of the 2022 ACM Conference on Human Factors in Computing Systems* (pp. 1–13). <https://doi.org/10.1145/3491102.3517502>
- Schegloff, E. A. (1997). Whose text? Whose context? *Discourse & Society*, 8(2), 165–187. <https://doi.org/10.1177/0957926597008002002>
- Verma, H., Mlynář, J., Schaer, R., Reichenbach, J., Jreige, M., Prior, J. O., & Evéquo, F. (2023). Rethinking the role of AI with physicians in oncology: Revealing physicians' perspectives on AI integration. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). <https://doi.org/10.1145/3544548.3581506>
- Bowker, G., Star, S. L., Gasser, L., & Turner, W. (Eds.). (2014). *Social science, technical systems, and cooperative work: Beyond the great divide*. Routledge. <https://doi.org/10.4324/9781315805849>



11 JUNE 2025 09.00 - 11.00

ID 314 - STS interventions in causal AI: the case of clinical prediction model validation

James Lowe, University of Exeter

Niccolò Tempini, University of Exeter

Keywords: Causal AI, Clinical prediction models, anticipatory governance, interventionist STS, validation

As STS researchers leading RRI practices in CHAI - Causality in Healthcare AI Hub - involving six UK universities, we are engaged in interventions in the governance of novel technoscience.

CHAI has the stated ambition of developing Causal AI methods that raise scientific and ethical standards together. As they allow scientists to explicitly model scientific knowledge into the machine learner, Causal AI models should be more sustainable, scalable and reliable. They also create the opportunity for directly addressing the algorithmic causes of health outcome inequities. For the opportunity to be seized, the CHAI Hub recognizes the need to adopt a model of anticipatory research governance centred on bodies convening different configurations of Hub affiliates. Through using social scientific co-creation methods, these bodies will ensure that expertise and stakeholder contributions are available to inform key decision-making and maximise social impact.

Anticipatory research governance requires qualitative research to be a key component of the Hub's strategy. Accordingly, we aim to intervene in scientific debates as they develop, and to use our attention to data practices and their wider social context to inform practical suggestions and plans for action. Our ambition is to espouse and enact an expansive definition of Causal AI, as not limited to embedding scientific theory and demographic considerations but including knowledge of social systems as well.

In this paper, we discuss how we are using qualitative research methods to directly contribute to a key frontier in causal AI development: the targeted validation of clinical prediction models. Clinical prediction models aid in diagnosis, prognosis, risk prediction and therefore in informing testing and treatment options for patients. Clinical prediction models can have external validity problems due to the inescapable differences between the context of model training and the context of model application.

Targeted validation is an approach to developing and rolling out clinical prediction models that has been argued for by health statisticians (Sperrin et al 2022). It calls for developing methods to allow local modification and validation of models. This should help ensure maintenance of performance across implementation sites, thereby improving the distribution of outcomes across different populations. Translation of the model beyond the initial context requires assessment of a potential 'validation gap' then the modification of the model to fit the characteristics of the target population and setting, and its consequent validation.

As qualitative researchers we contribute to the endeavour of identifying and responding to drivers of validation gaps by pushing the agenda of targeted validation beyond concerns with data sources and demographic differences identified in the literature and by developing social scientific techniques for addressing anything relevant that cannot be directly modelled. This will encompass the broader context of the delivery of healthcare, including socio-technical process analysis. At stake is the scalability and systemic feasibility of the whole approach.

In our paper we will explicate our research and engagement concerning targeted validation and will elaborate lessons for the anticipatory governance model CHAI is committed to.



11 JUNE 2025 09.00 - 11.00

ID 551 - Scalable Screenings: A Critical Analysis of Narratives Surrounding Mental Health Apps

Greta Von Albertini, University College London

Keywords: mental health, AI, personalisation, apps, narratives

Mental health smartphone applications (MH apps) are presented as urgent and scalable solutions to rising mental illness rates that burden individuals and societies worldwide, offering low-cost interventions, increased accessibility, reduced stigma, and flexibility. However, framings of a 'global mental health epidemic' neglect systemic inequalities and "underlying social causes of mental ill health" (Rose et al., 2021: 4), resulting in an erroneous sense of a 'universal cure' as desirable—or even possible. While the technologisation of mental health care shapes experiences of mental illness and therapy, it also raises critical questions of who bears responsibility for 'getting better'. Drawing on digital ethnographic fieldwork with an MH app company, my research examines how disruptions to existing care infrastructures are legitimised through narratives emphasising scalable and personalised care, alongside expectations that the future of therapy will inevitably involve AI. I investigate how such narratives draw on psychiatric and computational epistemologies and how these are enacted through conceptualisations of 'users' as quantifiable and diagnosable, where diagnostic screenings classify individuals as either healthy or unhealthy, and personalised data analysis determines appropriate therapy. I argue that these narratives depoliticise and commodify both care work and suffering while sidelining alternative approaches to mental health. This research challenges claims for the marketisation and automation of mental health care. It contributes to the sociology of expectations and literature on sociotechnical imaginaries, as well as the politics of scaling and personalisation in healthcare.

