

Stimulating cognitive engagement in hybrid decision-making: friction, reliance and biases

Workshop Proposal

Chiara NATALI^{a,1}, Brett FRISCHMANN^b and Federico CABITZA^{a,c}

^a *DISCo, University of Milano-Bicocca, viale Sarca 336, Milano, 20126, Italy*

^b *Charles Widger School of Law, Villanova University, 299 N. Spring Mill Rd. Villanova, Pennsylvania, United States*

^c *IRCCS Orthopedic Institute Galeazzi, via Galeazzi, 4, Milano, 20161, Italy*

Abstract. This workshop critically examines the trend of pursuing increasingly rapid and effortless interaction with AI, challenging the traditional view that human over-reliance on AI stems solely from inherent and unavoidable cognitive biases. Instead, we highlight the crucial role of designers and programmers in fostering user empowerment, skill enhancement, and responsibility. Our approach advocates for a thoughtful balance in Human-AI interaction, harmonizing operational efficiency with the necessity for effective, ethical human knowledge work. At the heart of our discourse is the notion of 'programmed inefficiencies' or 'frictional protocols' in AI systems. These are intentionally integrated to engage users cognitively, fostering interactions that are mindful, even if they might be slower. We welcome a diverse range of contributions, spanning innovative design principles that strike a balance between efficiency and cognitive engagement, to methodologies for assessing and reducing both over-reliance and under-reliance on AI systems. Our goal is to explore and develop strategies that encourage more thoughtful, informed interactions between humans and AI through what we termed 'Frictional AI'.

Keywords. Human-AI Interaction, Frictional AI, Decision Support Systems, Machine Learning, Interaction protocols, Usability

Name, affiliation, and short bio of each organizer

Prof. Federico Cabitza, BSc, MEng, PhD

Federico Cabitza is an Associate Professor at the University of Milano-Bicocca (Milan, Italy) where he teaches various classes, including human-computer interaction, interaction design, information systems and decision support. He is head of the Laboratory of Uncertainty Models, Decisions and Interactions (MUDILab) in the department of Informatics at the above-mentioned university and is director of the local node of the CINI

¹Corresponding author: chiara.natali@unimib.it

national laboratory “Computer Science and Society.” Since 2016 he has been intensively collaborating with several hospitals, among which the IRCCS Hospital Galeazzi and Sant’Ambrogio in Milano (Italy), with which he is formally affiliated as senior researcher and where he co-founded the Medical AI laboratory. He is associate editor of the International Journal of Medical Informatics (Elsevier ISSN: 1386-5056) and a member of several editorial boards, including the Machine Learning and Knowledge Extraction journal (ISSN 2504-4990), the Journal of Medical Artificial Intelligence (ISSN 2617-2496), the Journal of Cross-disciplinary Research in Computational Law (CRCL - ISSN 2736-4321) and Mondo Digitale, the official AICA journal. In 2008, he held a workshop with Prof. Batini, Pasi and Schettini on “Quality of Data, Textual Information and Images” at the 27th International Conference on Conceptual Modeling (ER2008) in Barcelona, Spain. He has co-chaired International workshops (on Data Visualization in Healthcare and knowledge IT artifacts), conference tracks (on Socio-technical design at ECIS), conference programs (for the Italian Chapter of AIS, Healthinf 2020 and 2023, part of BIOSTEC conference) and special issues on impacted Journals (i.e., the Health Informatics Journal by SAGE, CSCW journal by Springer, and Program, By Emerald). He held a number of keynote speeches, among which at IFIP CD-MAKE 2020, BIOSTEC 2021 and ISOC 2022. His research interests are the design and evaluation of artificial intelligence systems to support decision making, especially in health care and law, and the impact of these technologies on the organizations that adopt them and user experience and work processes. To date, he has published more than 150 research publications in international conference proceedings, edited books and high-impact scientific journals and is listed among the world’s most influential scientists, according to Stanford’s Top 2% Scientists list. He is the author with Prof. Luciano Floridi (former University of Oxford UK, now University of Bologna, Italy and Yale University, CT, USA) of the book “Artificial Intelligence, the use of the new machines” published by Bompiani, one of the leading Italian publishing houses (founded in 1929 in Milan).

Prof. Brett Frischmann, PhD

Brett Frischmann is The Charles Widger Endowed University Professor in Law, Business, and Economics at Villanova University (Philadelphia, PA, USA). In addition, Frischmann serves as an Affiliate Scholar at the Center for Internet and Society at Stanford Law School, an Affiliated Faculty member of the Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis at Indiana University, and as a Trustee for the Nexa Center for Internet & Society in Torino, Italy. Frischmann also served as the Microsoft Visiting Professor of Information and Technology Policy at Princeton University’s Center for Information and Technology Policy. Frischmann’s teaching expertise includes intellectual property, Internet law, privacy, and technology policy, with research interests spanning various disciplines and topics. His interdisciplinary research on the relationships between infrastructural resources, governance, commons, and spillovers includes articles in law reviews, peer reviewed journals, conference proceedings, and other interdisciplinary research venues as well as a series of foundational books, including “Infrastructure: The Social Value of Shared Resources” (Oxford UP 2012), “Governing Knowledge Commons” (Oxford UP 2014, with Katherine Strandburg and Michael Madison), “Governing Medical Knowledge Commons” (Cambridge UP 2017, with Strandburg and Madison), “Governing Privacy in Knowledge Commons” (Cambridge UP 2021,

with Madelyn Sanfilippo and Strandburg) and “Governing Smart Cities as Knowledge Commons” (Cambridge 2023, with Madison and Sanfilippo). In 2018, Cambridge University Press published “Re-Engineering Humanity,” which Frischmann co-authored with Evan Selinger. This wide-reaching, interdisciplinary book examines what’s happening to our lives as society embraces big data, predictive analytics, and smart environments. They explain how the goal of designing programmable worlds goes hand in hand with engineering predictable and programmable people. Detailing new frameworks, provocative case studies, and thought experiments, Frischmann and Selinger reveal hidden connections between fitness trackers, electronic contracts, social media platforms, robotic companions, fake news, autonomous cars, and more. Among other things, they make a strong case for pro-social friction-in-design and against imperatives that favor seamless and frictionless design.

Chiara Natali, BSc, MA

Chiara Natali is a PhD student in Informatics at University of Milano-Bicocca with a background in Political Philosophy. Her interdisciplinary understanding of the most pressing ethical challenges of human-AI interaction and eXplainable AI underscores her research on the socio-technical shortcomings behind biased decision-making. Her research work on decision support systems that leverage human intuition received the “Best Doctoral Consortium Paper” accolade at CHIItaly23. She is the Teaching Assistant of the Human-Computer Interaction course at the BSc in Computer Science, PhD representative at the faculty board, and Science Ambassador within the Gender Equality Plan at University of Milano-Bicocca. A strong believer in scientific divulgation since her participation to TEDxBocconiU, her dissemination activities involved two events for Milano Digital Week (University of Milano-Bicocca, European Institute of Design - IED), three guest lectures (at La Sapienza University, University of Pavia, and IED), as well as holding and organizing tutorials at conferences (HHAI’23, INTERACT’23) on “How to Assess Human Reliance on Artificial Intelligence in Hybrid Decision-Making”.

Topics and issues

The rapid growth of generative and conversational AI systems, capable of creating complex content and engaging in persuasive interactions, has raised critical issues in Human-AI Interaction. These systems are designed to align closely with human thinking, aiming to build trust and reduce aversion. However, this may lead to an increased reliance on AI advice, spurred by user-friendly interfaces that reduce cognitive effort.

Instead of attributing the phenomenon of human over-reliance on AI systems solely to cognitive biases and human limitations, this workshop emphasizes the central role of designers and programmers in fostering user agency, skill, and responsibility. We challenge the norm of prioritizing seamless Human-AI interaction, advocating instead for a responsible approach that balances efficiency with the efficacy and integrity of human knowledge work. In line with slow design principles, we promote thoughtful user engagement, deliberate action and skill retention through the concept of ‘Frictional AI’, which leverages ‘programmed inefficiencies’ or ‘frictional protocols’ in AI systems.

Contributions are invited on topics including, but not limited to:

- **Novel Design Principles for Cognitive Engagement:** Proposing a balance between efficiency and cognitive engagement in AI design.
- **Measurement and Mitigation Strategies for Over-reliance and Under-reliance:** Introducing frameworks to evaluate AI's impact on human judgment through novel metrics and/or proposals to mitigate risks like automation bias and deskilling.
- **Calibration of Appropriate Trust in AI:** Investigating the promotion of appropriate levels of trust in AI models by users, for example scrutinizing how transparency can both aid and hinder trust calibration.
- **Governance solutions for Cognitively Engaging AI Design:** Discussing policy and governance approaches to promote cognitive engagement and frictional principles in AI, fostering responsible and ethical AI development.
- **Applications and Case Studies:** Works documenting and demonstrating practical applications of seamful/frictional principles in various settings, supported by user studies and/or open-source tools.

By emphasizing the need for cognitive engagement and responsible design in AI systems, this workshop caters to both scholar and practitioners to promote more intentional, ethical, and effective use of AI across various domains to enhance human decision-making capabilities.

Table 1. Definition of all possible decision- and reliance-patterns between human decision makers and their AI system, focusing on the phenomena of automation bias and detrimental algorithmic aversion. In the first three columns, 0 denotes an incorrect decision, and 1 a correct decision.

Human judgment (H)	judg- ment (H)	AI support ² (AI)	Final decision (D)	Reliance pattern	Biases and Effects
0		0	0	detrimental reliance	automation complacency
0		0	1	beneficial under-reliance	extreme algorithmic aversion
0		1	0	detrimental self-reliance	detrimental algorithmic aversion
0		1	1	beneficial over-reliance	algorithm appreciation
1		0	0	detrimental over-reliance	automation bias
1		0	1	beneficial self-reliance	beneficial algorithmic aversion
1		1	0	detrimental under-reliance	extreme algorithmic aversion
1		1	1	beneficial reliance	confirmation bias (in later cases)

Intended workshop duration

This workshop is conceived as a half-day event.

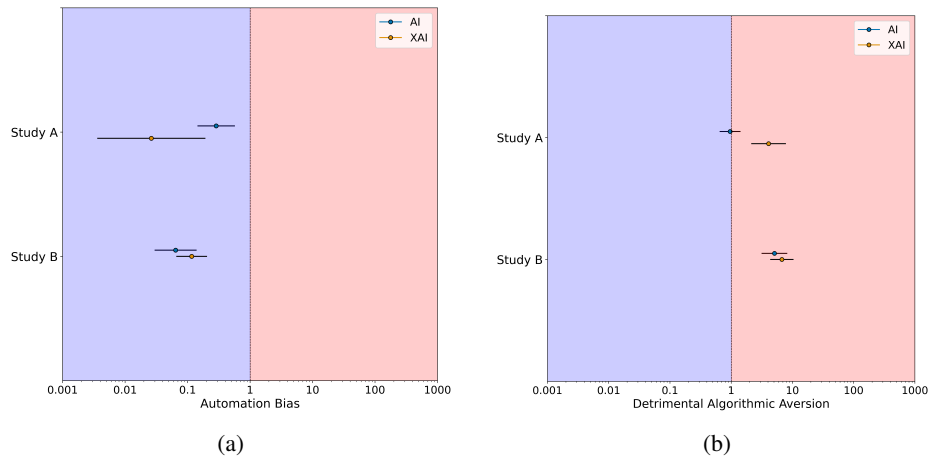


Figure 1. An example of Automation Bias (a) and Detrimental Algorithmic Aversion (b) odds ratios as defined in Table 1, that is, (a) whenever AI support may detrimentally lead individuals to discard their correct initial decisions for incorrect final decisions, and (b) whenever AI support is not able to lead individuals to discard their incorrect initial decisions despite correct AI advice.

Discussion of closely related workshops and tutorials

This workshop is intended to be the first of its kind in its discussion of Frictional AI, a novel concept that draws from several intuitions and reflections from the Hybrid Human-AI Interaction research community. As such it is inspired by, and aims to expand on, previous contributions of the authors in workshops dedicated to Human-Centered AI which were organized at flagship conferences (see e.g. HCAI@NEURIPS 2021³ and Realizing AI in Healthcare at CHI 2021⁴), including the HHAI'23 conference, with the tutorial "How to Assess Human Reliance on Artificial Intelligence in Hybrid Decision-Making"⁵, also held at INTERACT'23⁶ and to be replicated at HCII'24. In these workshops, we proposed a brief, conceptual illustration of the consequentialist and interactionist approaches to evaluating the impact and effect of AI support in hybrid decision making.

Workshop format

The workshop will be structured in two parts and a final discussion. In the first part, the organizers will provide a short introduction on their specific understandings on Human-AI Reliance and Frictional AI, in the form of a traditional frontal lecture with slides. The second part will be constituted by the presentation of workshop papers submitted by the participants; a final round-table and wrap-up discussion among the participants and organizers will conclude the work of the workshop about what it means to have

³<https://sites.google.com/view/hcai-human-centered-ai-neurips>

⁴<http://franciscoconunes.me/RealizingAIinHealthcareWS/>

⁵<https://mudilab.wixsite.com/ai-reliance-tutorial>

⁶<https://interact2023.org/courses/>

a quali-quantitative assessment of human reliance on AI, the promotion of cognitive engagement and bias avoidance, as to build a common understanding of these issues within the framework of Frictional AI for future research within the HHAI community.

Target audience size

15 to 20 people.

Promotional Activities

Following the publication of the Workshop website, we will attract people via a Call for Abstracts which we will diffuse by social media (Twitter, LinkedIn) and mailing lists (e.g. CINI – the Italian Inter-University Consortium of Informatics; GRIN – the Informatics Italian Research Group; Cyberprof; Privacy listsrvs), as well as personal invitations to scholars that published works aligned to the concept of Frictional AI and contacting like-minded research groups and research initiatives.

Programme Committee

The following alphabetical list of scholars represents the provisional Programme Committee, to be expanded following the confirmation of the workshop:

- Noah Apthorpe (Colgate University, USA, Computer Science)
- Andrea Campagner (IRCCS Galeazzi Sant’Ambrogio Hospital, Italy, Artificial Intelligence)
- Davide Ciucci (University of Milano-Bicocca, Italy, Computer Science)
- Vincenzo Crupi (University of Turin, Italy, Philosophy)
- Diletta Huyskes (University of Milan, Italy, Sociology)
- Enea Parimbelli (University of Pavia, Italy, Engineering)
- Sarah Michele Rajtmajer (Pennsylvania State University, USA, Computer Science)
- Evan Selinger (Rochester Institute of Technology, USA, Philosophy)
- Yan Shvartzshnaider (York University, UK, Computer Science)
- Alberto Termine (IDSIA USI-SUPSI, Switzerland, Artificial Intelligence)

Special Requirements

The workshop should be held in the afternoon to facilitate remote participation from the USA. It also requires necessary equipment to ensure remote participation is possible (computer, webcam, microphone).

Supporting material

Further readings

Cabitza, F., Campagner, A., Angius, R., Natali, C., Reverberi, C. (2023). AI Shall Have No Dominion: on How to Measure Technology Dominance in AI-supported Human decision-making. In CHI '23: the Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, April 23–28, 2023, Hamburg, Germany.

Frischmann, B., & Selinger, E. (2018). *Re-engineering humanity*. Cambridge University Press.

Cabitza, F., Campagner, A., Ronzio, L., Cameli, M., Mandoli, G. E., Pastore, M. C., ... & Gamboa, H. (2023). Rams, hounds and white boxes: Investigating human–AI collaboration protocols in medical diagnosis. *Artificial Intelligence in Medicine*, 138, 102506.

Natali, C. (2023). Per Aspera ad Astra, or Flourishing via Friction: Stimulating Cognitive Activation by Design through Frictional Decision Support Systems. In *CEUR WORKSHOP PROCEEDINGS* (Vol. 3481, pp. 15-19).

Cabitza, F., Campagner, A., Natali, C., Parimbelli, E., Ronzio, L., & Cameli, M. (2023). Painting the black box white: experimental findings from applying XAI to an ECG reading setting. *Machine Learning and Knowledge Extraction*, 5(1), 269-286.

Cabitza, F., & Natali, C. (2022). Open, multiple, adjunct. Decision support at the time of relational AI. In *HHAI2022: Augmenting Human Intellect* (pp. 243-245). IOS Press.

Frischmann, B. M., & Benesch, S. (2022). Friction-In-Design Regulation as 21St Century Time, Place and Manner Restriction. *Place and Manner Restriction* (August 1, 2022).

Cabitza, F. (2021). Cobra AI: Exploring Some Unintended Consequences. *Machines We Trust: Perspectives on Dependable AI*, 87.

Ohm, P. & Frischmann, B. M. (Forthcoming). *Governance Seams*. *Harvard Journal of Law and Technology*.

Web resources

DSS Quality Assessment Tool: <https://dss-quality-assessment.vercel.app>