

Stimulating Cognitive Engagement in Hybrid Decision-Making: Friction, Reliance and Biases (preface)

Chiara Natali^{1,*}, Brett Frischmann² and Federico Cabitza^{1,3}

¹University of Milano-Bicocca, Viale Sarca 336, Milan, Italy

²Villanova University, 299 N. Spring Mill Rd. Villanova, Pennsylvania, United States

³IRCCS Galeazzi Sant'Ambrogio Hospital, Via Cristina Belgioioso 173, Milan, Italy

Abstract

This workshop critically examined the trend toward rapid and seamless human–AI interactions and considered alternative forms of prosocial engagement. We focused on the role of designers and developers in fostering user empowerment, skill development, and appropriate reliance on AI for responsible decision-making. Our discussions centered on friction-in-design and the core concepts of 'programmed inefficiencies' and 'frictional protocols' that involve design elements intentionally included to promote cognitive engagement and thoughtful interaction with AI, even if they might be slower. The workshop featured contributions on design principles that balance efficiency with engagement, methods for revealing and reducing biases in explainable AI systems, and considerations for a meaningful future with AI. In this first edition, this workshop set the stage for future research and community-building efforts around 'Frictional AI' to encourage more informed and reflective human-AI interactions.

Keywords

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

1. Introduction

We are pleased to present the proceedings of the inaugural Frictional AI Workshop, which took place on June 11th, 2024 at HHAI2024 (Malmö, Sweden) as a half-day event. This workshop marked a first milestone in the exploration of Frictional AI, a novel concept aimed at redefining the dynamics of Human-AI Interaction. By challenging the prevailing trends that favor seamless and rapid interaction with AI systems, the workshop sought to introduce and examine 'frictional protocols' [1]—deliberate design choices that slow down interactions to foster greater cognitive engagement and more thoughtful decision-making.

HHAI-WS 2024: Workshops at the Third International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 10–14, 2024, Malmö, Sweden

*Corresponding author.

✉ chiara.natali@unimib.it (C. Natali); brett.frischmann@law.villanova.edu (B. Frischmann); federico.cabitza@unimib.it (F. Cabitza)

🌐 <https://sites.google.com/view/chiaranatali/> (C. Natali); <http://www.brettfrischmann.com> (B. Frischmann); <https://www.federicocabitza.net> (F. Cabitza)

🆔 0000-0002-5171-5239 (C. Natali); 0000-0002-7425-8931 (B. Frischmann); 0000-0002-4065-3415 (F. Cabitza)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

The workshop brought together a diverse group of scholars, practitioners, and researchers who critically examined the role of AI designers and developers in shaping user reliance on AI systems. Moving beyond the conventional attribution of over-reliance to inherent cognitive biases, the discussions highlighted how intentional design can either exacerbate or mitigate such biases, ultimately influencing the quality of human knowledge work and decision-making.

The workshop's contributions can be broadly categorized into two core themes, encompassing both theory and practice: the theoretical exploration of biases in Human-AI interaction and the presentation of practical design applications.

The first theme focused on a theoretical examination of cognitive biases and their interaction with AI systems. The contributions involved reflections on the psychological and philosophical factors influencing human reliance on AI, as well as proposing philosophical accounts over a meaningful future with AI. This theoretical exploration provided a foundation for understanding how our knowledge of human biases and reflections on the future of Human-AI interaction can be leveraged to improve decision-making quality through more reflective and deliberate interactions with AI.

The second core theme centered around the practical application of frictional design principles in real-world settings. Case studies and design frameworks were presented, illustrating how frictional protocols can be integrated into various AI systems to balance efficiency with cognitive engagement. Participants shared insights into how these principles could be applied across different domains, from seamless design for human-AI creative systems to adding friction to human-robot interaction, decision support systems for social media content regulation, and AI-mediated communication in medicine. The discussions in this theme provided clear examples of how frictional design can help prevent automation bias, encourage skill retention, and promote ethical AI development and use.

2. Organization

2.1. Workshop Chairs

- Chiara Natali (University of Milano-Bicocca, Italy)
- Brett M. Frischmann (Villanova University, USA)
- Federico Cabitza (University of Milano-Bicocca, IRCCS Galeazzi Sant'Ambrogio Hospital, Italy)

2.2. Programme Committee

The Programme Committee comprised a multidisciplinary team of experts from fields including Computer Science, Human-Centered Computing, Human-Computer Interaction, Psychology, Philosophy, Sociology, and Artificial Intelligence. Their collective expertise was instrumental in ensuring the rigorous evaluation of workshop submissions.

- Noah Apthorpe (Colgate University, USA, Computer Science)
- Niels van Berkel (Aalborg University, Denmark, Human-Centred Computing)
- Andrea Campagner (IRCCS Galeazzi Sant'Ambrogio Hospital, Italy, Artificial Intelligence)

- Marta E. Cecchinato (Northumbria University, UK, Human-Computer Interaction)
- Paolo Cherubini (University of Pavia, Italy, Psychology)
- Lewis L. Chuang (Chemnitz University of Technology, Germany, Neuroscience)
- Davide Ciucci (University of Milano-Bicocca, Italy, Computer Science)
- Vincenzo Crupi (University of Turin, Italy, Philosophy)
- Diletta Huyskes (University of Milan, Italy, Sociology)
- Jo Iacovides (University of York, UK, Human-Computer Interaction)
- Sarah Inman (Google, USA, Human-Centered Design)
- Tomáš Kliegr (Prague University of Economics, Czechia, Informatics)
- Tim Miller (University of Queensland, Australia, Artificial Intelligence)
- Mohammad Naiseh (Bournemouth University, England, Artificial Intelligence)
- Enea Parimbelli (University of Pavia, Italy, Engineering)
- Sarah Michele Rajtmajer (Pennsylvania State University, USA, Computer Science)
- Carlo Reverberi (University of Milano-Bicocca, Italy, Psychology)
- David Ribes (University of Washington, USA, Sociology)
- Scott Robbins (University of Bonn, Germany, Ethics of AI)
- Evan Selinger (Rochester Institute of Technology, USA, Philosophy)
- Yan Shvartzshnaider (York University, Canada, Computer Science)
- Alberto Termine (IDSIA USI-SUPSI, Switzerland, Artificial Intelligence)

3. Summary of the workshop

The workshop included 8 accepted submissions, with authors from institutions from Italy, United States of America, Germany Portugal and Sweden.

The submissions were grouped according to their overarching themes, identifying two presentation sessions: Human-AI Collaboration and Biases and Frictional AI applications.

Each session included a reflection roundtable, where all the paper authors discussed the similarities and differences of their approaches and answered questions from the audience. Finally, we discussed future work to build the frictional AI community.

Introductory talks

- Brett M. FRISCHMANN, Villanova University (USA) - "An Interdisciplinary Research Agenda for Prosocial Friction-in-Design"
- Chiara NATALI, University of Milan-Bicocca (Italy) - "Frictional AI: Topics and Issues"

Brett M. Frischmann's talk, "An Interdisciplinary Research Agenda for Prosocial Friction-in-Design," drew from his 2018 book *Re-Engineering Humanity* with Evan Selinger [2] and subsequent research on friction-in-design [3, 4]. He addressed the root of humanity's techno-social dilemma: the prevailing economic, social, and political logics that drive the design of AI systems toward goals like maximizing efficiency, minimizing transaction costs, and eliminating friction [3]. Frischmann argued that these design principles, which prioritize speed,

scale, and seamlessness, often undermine human autonomy and social welfare. To counter these tendencies, he called for prosocial "friction-in-design" principles and regulations that challenge the conventional wisdom perpetuating these logics. His proposed strategies include intentionally engineering friction, such as transaction costs and inefficiencies, into AI systems to resist the dominance of efficiency and productivity logics and to promote human flourishing through the exercise and development of human capabilities [2]. Chiara Natali followed with "Frictional AI: Topics and Issues," providing a comprehensive overview of the key areas and challenges in applying frictional design to AI systems drawing parallels with slow design [5], microboundaries [6], desirable and programmed inefficiencies [7, 8] for constructive distrust [9] and debiasing strategies against over-confidence [10, 11] and anchoring bias [12]. This requires new methodologies to assess over- and under-reliance [13], such as the *Human-AI Interaction Assessment* tool. Together, these talks set the stage for a deeper exploration of how frictional design can be strategically used to shape the social and ethical impacts of AI.

First session: Human-AI Collaboration and Biases

- Regina DE BRITO DUARTE and Joana CAMPOS, INESC-ID, Instituto Superior Tecnico (Portugal) - "Looking for cognitive bias in Human-AI decision-making"
- Sebastiano MORUZZI, Filippo FERRARI and Filippo RISCICA LIZZO, University of Bologna (Italy) - "Biases, Epistemic Filters, and Explainable Artificial Intelligence"
- Christopher D. QUINTANA, Georg THEINER, Villanova University (USA) - "Make Friends, Not Tools: Designing AI for Technoamicitia"
- Scott ROBBINS, University of Bonn (Germany) - "Beyond Regulation: How We Can Craft a Meaningful Future with AI"

The *Human-AI Collaboration and Biases* session examines different aspects of bias and interaction in human engagement with AI systems, emphasizing the need to rethink design and user practices to promote thoughtful and meaningful AI deployment. Regina de Brito Duarte and Joana Campos examine cognitive biases in AI-assisted decision-making, advocating for balanced friction to avoid both over-reliance and undue skepticism towards AI recommendations. Sebastiano Moruzzi, Filippo Ferrari, and Filippo Riscica discuss how "epistemic filters" impact the outputs and user interactions with XAI and Generative AI, and how understanding and adjusting them can address technical and cognitive biases. Christopher D. Quintana and Georg Theiner propose "technoamicitia," a design approach that goes beyond traditional usability metrics to foster deeper human engagement with AI: their approach aims to support psychological and moral development, and thus counter the prevailing view of AI as mere tools for efficiency and productivity. Scott Robbins builds upon the concept of friction by challenging the conventional focus on regulation and design controls as sole means of achieving ethical AI deployment; he suggests that norms around the intentional use or restraint of AI can help preserve human autonomy and ensure that certain meaningful tasks remain within human control.

Second session: Frictional AI Applications

- Caterina FREGOSI, Federico CABITZA, University of Milan-Bicocca (Italy) - "A frictional design approach: towards Judicial AI and its possible applications"

- Ingar BRINCK, Samantha STEDTLER and Valentina FANTASIA, Lund University (Sweden) - "Exploring Frictional Design in Human-Robot Interaction: Delayed Movement in a Turn-taking Game"
- Sarah INMAN and Sarah D'ANGELO, Google (USA) - "Enabling Creative Human-AI Systems with Seamful Design" (*not included in the proceedings*)
- Evan SELINGER, Rochester Institute of Technology (USA) - "Balancing Empathy and Accountability: Exploring Friction-In-Design For AI-Mediated Doctor-Patient Communication"

The *Frictional AI Applications* session highlights diverse approaches to incorporating intentional friction in AI design to promote critical thinking, creativity, and ethical engagement. Caterina Fregosi and Federico Cabitza present "Judicial AI," a decision support system that offers two contrasting explanations to foster critical thinking and reduce automation bias. They explore how complex decision pathways can enhance user autonomy. Ingar Brinck, Samantha Stedtler, and Valentina Fantasia examine frictional design in human-robot interactions and demonstrate how deliberate delays in a turn-taking game can enhance cognitive engagement and foster deeper interaction with social robots. Sarah Inman and Sarah D'Angelo propose applying "seamful design" in software engineering to support creative problem-solving: they emphasize the value of exposing hidden processes to maintain control and foster creativity beyond mere productivity. Evan Selinger suggests using generative AI to enhance empathetic content in doctor-patient communication, addressing the issue of doctors often sounding robotic due to systemic pressures. To ensure this technology is used ethically and maintains trust, he advocates for incorporating friction, such as transparency measures and manual revisions, and establishing governance procedures to hold doctors accountable for how they integrate AI-generated content into their messages.

4. Conclusion and Remarks

The concept of Frictional AI draws heavily on the idea that some level of friction, or 'seamfulness,' is essential to prevent overreliance on AI and to maintain human agency in decision-making processes. As Frischmann and Selinger [2] argued in *Re-engineering Humanity*, tolerating some friction in our interactions with technology is vital for sustaining environments that support human flourishing.

The Frictional AI Workshop has laid the groundwork for future research and collaboration on this new paradigm in Human-AI Interaction—one that values cognitive engagement and ethical responsibility as much as it does efficiency and performance.

Looking ahead, we are confident that the contributions contained in these proceedings will serve as a valuable resource for scholars and practitioners alike, providing both theoretical frameworks and practical guidance for integrating Frictional AI into a wide range of applications.

Acknowledgments

We extend our sincere gratitude to all the participants, speakers, and the HHAI conference organizers who contributed to the success of this workshop. Special thanks go to the members

of the Programme Committee for their expertise and commitment.

C. Natali gratefully acknowledges the PhD grant awarded by the Fondazione Fratelli Confalonieri, which has been instrumental in facilitating her research pursuits.

F. Cabitza acknowledges funding support provided by the Italian project PRIN PNRR 2022 InXAID - Interaction with eXplainable Artificial Intelligence in (medical) Decision making. CUP: H53D23008090001 funded by the European Union - Next Generation EU.

References

- [1] F. Cabitza, C. Natali, L. Famigliani, A. Campagner, V. Caccavella, E. Gallazzi, Never tell me the odds: Investigating pro-hoc explanations in medical decision making, *Artificial Intelligence in Medicine* 150 (2024) 102819.
- [2] B. Frischmann, E. Selinger, *Re-engineering humanity*, Cambridge University Press, 2018.
- [3] B. Frischmann, S. Benesch, Friction-in-design regulation as 21st century time, place, and manner restriction, *Yale JL & Tech.* 25 (2023) 376.
- [4] B. Frischmann, P. Ohm, Governance seams, *Harvard Journal of Law & Technology* Volume 37 (2023).
- [5] B. Grosse-Hering, J. Mason, D. Aliakseyeu, C. Bakker, P. Desmet, Slow design for meaningful interactions, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 3431–3440.
- [6] A. L. Cox, S. J. Gould, M. E. Cecchinato, I. Iacovides, I. Renfree, Design frictions for mindful interactions: The case for microboundaries, in: *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*, 2016, pp. 1389–1397.
- [7] P. Ohm, J. Frankle, Desirable inefficiency, *Fla. L. Rev.* 70 (2018) 777.
- [8] F. Cabitza, A. Campagner, D. Ciucci, A. Seveso, Programmed inefficiencies in dss-supported human decision making, in: *Modeling Decisions for Artificial Intelligence: 16th International Conference, MDAI 2019, Milan, Italy, September 4–6, 2019, Proceedings 16*, Springer, 2019, pp. 201–212.
- [9] M. Hildebrandt, Privacy as protection of the incomputable self: From agnostic to agonistic machine learning, *Theoretical Inquiries in Law* 20 (2019) 83–121.
- [10] T. Kliegr, Š. Bahník, J. Fürnkranz, A review of possible effects of cognitive biases on interpretation of rule-based machine learning models, *Artificial Intelligence* 295 (2021) 103458.
- [11] A. Bertrand, R. Belloum, J. R. Eagan, W. Maxwell, How cognitive biases affect xai-assisted decision-making: A systematic review, in: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 2022, pp. 78–91.
- [12] A. K. P. Bach, T. M. Nørgaard, J. C. Brok, N. Van Berkel, “if i had all the time in the world”: Ophthalmologists’ perceptions of anchoring bias mitigation in clinical ai support, in: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–14.
- [13] F. Cabitza, A. Campagner, R. Angius, C. Natali, C. Reverberi, 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, 2023, to be published.

A. Online Resources

- Workshop website,
- Human-AI Interaction Assessment Tool.