Tim Räz

Researcher in Responsible AI, Bern, Switzerland.

tim.raez@posteo.de

Website | ORCID | Google Scholar | GitHub

---

1 Highlights

30+ publications; 700+ total citations (Google Scholar, 10/2026); h-index 16 (Google Scholar, 10/2026).

2 Academic Positions

2025–2026 Senior Researcher, Project *Science in Silico 2.0 – How the Combination of Computer Simulation and Machine Learning Transforms Science*, Institute of Philosophy, University of Bern, funded by Fritz Thyssen Foundation.

2020–2025 Postdoctoral Researcher, Project *Improving Interpretability. Philosophy of Science Perspectives on Machine Learning*, Institute of Philosophy, University of Bern, funded by SNSF.

2020 (Jul.-Sep.) Postdoctoral Researcher, Project *Socially acceptable AI and fairness trade-offs in predictive analytics*, Institute of Biomedical Ethics and History of Medicine, University of Zürich, funded by SNSF.

2019–2020 Postdoctoral Researcher, Project *Understanding Deep Learning*, Institute of Philosophy, University of Bern, funded by cogito foundation.

2017–2019 Lecturer (part-time), Institute of Philosophy, University of Bern.

2014–2017 Postdoctoral Researcher, Project *Agency and (Quantum) Physics*, Department of Philosophy, University of Konstanz, funded by TWCF.

2010–2013 Research and Teaching Assistant, Department of Philosophy, University of Lausanne (funded by SNSF).

2007 (spring) Research and Teaching Assistant, Institute of Philosophy, University of Bern.

3 Education

01/2019 MSc in Mathematics, University of Bern. Supervisor: Zoltan Balogh.

12/2013 PhD in Philosophy, University of Lausanne. Supervisor: Michael Esfeld.

11/2006 MSc in Philosophy, University of Bern. Supervisor: Gerd Grasshoff.

4 Publications

4.1 AI, Machine Learning & Computer Science (Conferences & Journals)

2026 Räz, T.: The ramifications of LLM watermarks for authorship and groups. **AI and Ethics** 6 (540).

Räz, T.: Transparent and Fair Profiling in Employment Services: Evidence from Switzerland. **Data & Policy** 8:e30.

2024 Räz, T.: Reliability Gaps Between Groups in COMPAS Dataset. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, **FAccT**, 113–126.

Räz, T.: Gerrymandering Individual Fairness. **Artificial Intelligence** 326: 104035.

2023 Räz, T.: Methods for Identifying Emergent Concepts in Deep Neural Networks. **Patterns** 4(6).

2022 Hertweck, C. and Räz, T.: Gradual (In)Compatibility of Fairness Criteria. Proceedings of the 36th AAAI Conference on Artificial Intelligence, **AAAI**, 36(11): 11926–11934.

2021 Räz, T.: Group Fairness: Independence Revisited. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, **FAccT**: 129–137.

4.2 Philosophy & Interdisciplinary (Journals)

2025 Räz, T., Pahud de Mortanges, A., Reyes, M.: Explainable AI in medicine: challenges of integrating XAI into the future clinical routine. **Frontiers in Radiology**, Vol. 5.

2024 Räz, T.: ML interpretability: Simple isn't easy. **Studies in History and Philosophy of Science** 103: 159–167.

Räz, T. and Beisbart, C.: The Importance of Understanding Deep Learning. **Erkenntnis** 89: 1823–1840.

2023 Jebeile, J., Lam, V., Maszjak, M., Räz, T.: Machine Learning and the Quest for Objectivity in Climate Model Parameterizations. **Climatic Change** 176: 101.

2022 Räz, T.: COMPAS: Zu einer wegweisenden Debatte über algorithmische Risikobeurteilung. **Forens. Psychol. Psychiatr. Krim.** 16: 300–306.

Räz, T.: Understanding Risk With FOTRES? **AI and Ethics** 3(4): 1153–1167.

Beisbart, C. and Räz, T.: Philosophy of Science at Sea: Clarifying the Interpretability of Machine Learning. **Philosophy Compass** 17(6): e12830.

Räz, T.: Understanding Deep Learning With Statistical Relevance. **Philosophy of Science** 89(1): 20–41.

2021 Jebeile, J., Lam, V. and Räz, T.: Understanding Climate Change with Statistical Downscaling and Machine Learning. **Synthese** 199: 1877–1897.

2018 Räz, T.: Euler's Königsberg: The Explanatory Power of Mathematics. **European Journal for Philosophy of Science** 8: 331–46.

2017 Räz, T.: The Volterra Principle Generalized. **Philosophy of Science** 84(4): 737–60.

Räz, T.: The Silent Hexagon: Explaining Comb Structures. **Synthese** 194(5): 1703–1724.

2015 Räz, T. and Sauer, T.: Outline of a Dynamical Inferential Conception of the Application of Mathematics. **Studies in History and Philosophy of Modern Physics** 49: 57–72.

Räz, T.: Say My Name: An Objection to Ante Rem Structuralism. **Philosophia Mathematica** 23(1): 116–25.

2013 Räz, T.: On the Application of the Honeycomb Conjecture to the Bee's Honeycomb. **Philosophia Mathematica** 21(3): 351–60.

Scholl, R., and Räz, T.: Modeling causal structures. **European Journal for Philosophy of Science** 3(1): 115–32.

4.3 Philosophy & Interdisciplinary (Book Chapters, Preprints)

2025 Räz, T.: From Explanations to Interpretability and Back. In Juan Durán and Giorgia Pozzi (Eds.): **Philosophy of Science for Machine Learning: Core Issues, New Perspectives**, Synthese Library.

2023 Räz, T.: Inter-Rater Reliability is Individual Fairness. arXiv:2308.05458.

2017 Räz, T. and Sauer, T.: The collaboration between Marcel Grossmann and Albert Einstein as a case of the application of mathematics. **Proceedings of the Fourteenth Marcel Grossmann Meeting on General Relativity**, 3368–3371.

2016 Räz, T.: The Necessity of Learning for Agency. philsci-archive:11956.

Räz, T.: Gone Till November: A disagreement in Einstein scholarship. In Sauer, T. and Scholl, R. (Eds.): **The Philosophy of Historical Case Studies**, Boston Studies in Philosophy and History of Science, Springer.

Scholl, R. and Räz, T.: Towards a Methodology for Integrated History and Philosophy of Science. In Sauer, T. and Scholl, R. (Eds.): **The Philosophy of Historical Case Studies**, Boston Studies in Philosophy and History of Science, Springer.

2014 Räz, T. and Sauer, T.: Ein Zyklenmodell der Anwendung von Mathematik. In Ralf Krömer and Gregor Nickel (Eds.): **Siegener Beiträge zur Geschichte und Philosophie der Mathematik** 4.

2013 Räz, T.: Comment on "The Undeniable Effectiveness of Mathematics in the Special Sciences". In M. C. Galavotti, S. Hartmann, M. Weber, W. Gonzalez, D. Dieks, and T. Uebel (Eds.): **New Directions in the Philosophy of Science**, Springer.

5 Scientific Reviewing

AFME 2024 @ NeurIPS (PC), AI & Ethics, AI & Society, AIES 2022 (PC), Apidologie, British Journal for Philosophy of Science, Erkenntnis, European Journal for Philosophy of Science, EWAF 2024 (PC), FAccT 2022, 2023, 2024 (PC), Grazer Philosophische Studien, Journal for General Philosophy of Science, Minds & Machines, NeurIPS 2026 (Position Papers Track, PC), Patterns, Philosophia Mathematica, Philosophy of Science, Philosophy & Technology, Science of Nature (Naturwissenschaften), Swiss National Science Foundation, Studies in History and Philosophy of Science, Synthese, Dialectica (former member EC).

6 Teaching

2024 (spring) Co-lecturer, Seminar: *AI in Medicine: Ethical and Legal Issues*, University of Bern.

2023 (spring) Co-lecturer, Colloquium: *Philosophy of Science*, University of Bern.

2018 (fall) Co-lecturer, Proseminar: *Transformative Experience*, University of Bern.

2018 (spring) Lecturer, Proseminar: *Scientific Explanation*, University of Bern.

2017 (fall) Lecturer, Proseminar: *Philosophy of Mathematics*, University of Bern.

2015 (fall) Lecturer, Proseminar: *The Structure of Scientific Revolutions*, University of Konstanz.

2012 (fall) Co-lecturer, Seminar: *Philosophy of Science*, University of Lausanne.

2007 (spring) Lecturer, Proseminar: *Introduction to Classical Logic*, University of Bern.

7 Presentations (selected)

2025 *Transparent and Fair Profiling in Employment Services: Evidence from Switzerland.* Philosophy of AI reading group (Rune Nyrup), Aarhus, Denmark (via Zoom).

Transparent and Fair Profiling in Employment Services: Evidence from Switzerland. Swiss State Secretariat for Economic Affairs (SECO), working group "Arbeitsmarkt / Arbeitslosenversicherung", Bern, Switzerland, October 7.

2024 *Reliability Gaps Between Groups in COMPAS Dataset*, virtual presentation at the ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), Rio de Janeiro, Brazil, June 3.

Authorship and the Politics and Ethics of LLM Watermarks. Colloquium of the Chair of Practical Philosophy, ETH Zürich, Switzerland May 15.

2023 *Reliability and Fairness.* Working group of Chair of Political Philosophy, University of Zürich, May 26.

Reliability and Fairness. University of Tübingen, Germany, April 19.

2022 *Gradual (In)Compatibility of Fairness Criteria* (with Corinna Hertweck) at AAAI 2022, Virtual Event, Vancouver, Canada, February 25.

2021 *On Model Interpretability* at the MCMP colloquium (via Zoom), LMU Munich, Germany, December 8.

On Model Interpretability at the Workshop "Machine Learning Meets Philosophy of Science", University of Tübingen, Germany, November 11.

Group Fairness: Independence Revisited at the ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Virtual Event, Canada, March 1–10.

2020 *Algorithmic Decision Making and Social Justice. On the Morals of Predictive Modeling* (with Corinna Hertweck, via Zoom) at the 5th European COST Conference on Artificial Intelligence in Finance and Industry, ZHAW Zürich, Switzerland, September 3.

The Impact of Statistics and Machine Learning on Understanding in Climate Modeling (with Julie Jebeile and Vincent Lam, via Zoom) at the conference "Data Science in Climate and Climate Impact Research", ETH Zürich, Switzerland August 21.

2019 *Understanding of DNNs, Understanding with DNNs, and Climate Modeling* at the Workshop "Big data, machine learning, climate modelling & understanding", University of Bern, Switzerland, November 15.

2018 *Euler's Königsberg: The Explanatory Power of Mathematics* at the Workshop "Mathematics and Mathematization of Physical Theory: Historical and Conceptual Aspects", University of Mainz, Germany, June 22 to 23.

2016 *The Volterra Principle Generalized* (in German) at the Oberseminar Geschichte der Mathematik und der Naturwissenschaften, University of Mainz, Germany, February 2.

2015 *Euler's Königsberg: The Explanatory Power of Mathematics* at the Workshop "New Work on Explanation and Understanding", MCMP Munich, Germany, May 1.

2014 *Towards a Methodology for Integrated History and Philosophy of Science* (with Raphael Scholl, University of Bern) at Integrated History and Philosophy of Science – &HPS5, University of Vienna, Austria, June 26 to 28.

2013 *Applying an Account of the Applicability of Mathematics to the World* (with Tilman Sauer, CalTech and University of Bern) at the Foundations of Physics Conference in Munich, Germany, July 29 to 31.

2012 *Comment on Mark Colyvan: The Ins and Outs of Mathematical Explanation* at the Workshop "Causation, Dispositions and Probabilities in Physics and Biology" in Lausanne, Switzerland, November 22–24.

The Applicability of Mathematics at the Workshop "Metaphysics of science: objects, relations and structures" in Lausanne, Switzerland, October 12–13.

2011 *Mathematics and Causality* at the Workshop "Causality and Mechanisms in Philosophy of Science" in Bern, Switzerland, November 24–25.

Why do we Model? (with Raphael Scholl, University of Bern) at the European Philosophy of Science Association Conference in Athens, Greece, October 5–8.