Curriculum Vitae
Dr Rune Nyrup

AOS: Philosophy of Science and Technology
Interdisciplinary Ethics of Artificial Intelligence
AOC: Philosophy of Healthcare
Philosophy of Archaeology

Centre for Science Studies, Department of
Mathematics, Aarhus University
rune.nyrup@css.au.dk
https://sites.google.com/view/runenyrup

Academic Positions

2023–   Associate Professor, Centre for Science Studies, Department of Mathematics, Aarhus U.
2020–23   Senior Research Fellow, Leverhulme Centre for the Future of Intelligence, U. Cambridge.
2017–20   Postdoc, Leverhulme Centre for the Future of Intelligence, U. Cambridge.

Affiliate appointments

2023–   Associate Fellow, Leverhulme Centre for the future of Intelligence, U. Cambridge.
2017–23   Research Fellow, Department of History and Philosophy of Science, U. Cambridge
2017–21   Postdoctoral By-fellow, Churchill College, Cambridge

---

Education

2013–17   PhD in Philosophy. Supervisors: Nancy Cartwright and Julian Reiss. Thesis title: *Hypothesis Generation and Pursuit in Scientific Reasoning.* Durham U.
2012–13   MPhil in History, Philosophy & Sociology of Science, Technology & Medicine, Uni. Cambridge.
2010–13   MA in Philosophy, U. Copenhagen.
2007–10   BA in Philosophy, U. Copenhagen.

Publications

Papers in peer-reviewed journals

1. Chu, C., Donato-Wodger, S., Khan, S., Shi, T., Leslie, K., Rahimi, S.A., **Nyrup, R.** and Grenier, A. 2024. Strategies to Mitigate Age-Related Bias in Machine Learning: Scoping Review. *JMIR Aging* 7:e53564.
2. **Nyrup, R.** 2023. Trustworthy AI: a plea for modest anthropocentrism. *Asian Journal of Philosophy* 2:40.
3. Chu, C., Donato-Wodger, S., Khan, S., **Nyrup, R.**, Leslie, K., Lyn, A., Shi, T., Bianchi, A., Rahimi, S.A. and Grenier, A. 2023. Age-related bias and artificial intelligence: a scoping review. *Humanities and Social Sciences Communications* 10: article no. 510.
4. **Nyrup, R.** 2022. The Limits of Value Transparency in Machine Learning. *Philosophy of Science* 89: 1054-1064.
5. **Nyrup, R.** and Robinson, D. 2022. Explanatory Pragmatism: A Context-sensitive Framework for Explainable Medical AI. *Ethics and Information Technology* 24:13.
6. Chu, C.H., **Nyrup, R.**, Leslie, K., Shi, J., Bianchi, A., Lyn, A., McNichol, M., Khan, S., Rahimi, S. and Grenier, A. 2022. Digital Ageism: Challenges and Opportunities in Artificial Intelligence for Older Adults. *The Gerontologist* 62(7):947-955.
7. Chu, C., Leslie, K., Shi, J., Bianchi, A., Khan, S., Rahimi, S., **Nyrup, R.** and Grenier, A. 2022. Age-related bias in AI: scoping review protocol. *JMIR Research Protocols* 11(6):e33211.
8. **Nyrup, R.** 2021. From General Principles to Procedural Values: Responsible Digital Health Meets Public Health Ethics, *Frontiers in Digital Health* 3:690417.
9. Cave, S., Whittlestone, J., **Nyrup, R.**, O hEigeartaigh, S. and R. Calvo. 2021. Using AI Ethically to Tackle Covid-19. *British Medical Journal* 2021;372:n364.

10. Ronquillo, CE., Peltonen, L., Pruinelli, L., Chu, C.H., Bakken, S. Beduschi, A., Cato, K., Hardiker, N., Junger, A., Michalowski, M., **Nyrup, R.**, et al. 2021. Artificial intelligence in nursing: Priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative, *Journal of Advanced Nursing* 77(9):3707-3717.

11. **Nyrup, R.** 2020. Of Water Drops and Atomic Nuclei: Analogies and Pursuit Worthiness in Science, *British Journal for the Philosophy of Science* 71(3): 881-903.

12. Stanley, D. and **Nyrup, R.** 2020. Strategies in Abduction: Generating and Selecting Diagnostic Hypotheses, *Journal of Medicine and Philosophy* 45(2): 159-178.

13. Porat, T., **Nyrup, R.**, Calvo, R., Paudyal, P. and Ford, E. 2020. Public Health and Risk Communication during COVID-19 – Enhancing Psychological Needs to Promote Sustainable Behaviour Change, *Frontiers in Public Health* 8:573397.

14. Cave, S., **Nyrup, R.**, Vold, K. and Weller, A. 2019. Motivations and Risks of Machine Ethics, *Proceedings of the IEEE* 107(3): 562-574.

15. **Nyrup, R.** 2015. How Explanatory Reasoning Justifies Pursuit: A Peircean View of IBE, *Philosophy of Science* 82(5): 749-760.

Peer-reviewed conference proceedings

16. Petti, U., **Nyrup, R.**, Skopek, J. and Korhonen, A. 2023. Ethical considerations in the early detection of Alzheimer's disease using speech and AI. *FAccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency,* 1062–1075.

17. Bermudez, J.P., **Nyrup, R.**, Deterding, S., Moradbakhti, L. Mougenot, C., You, F. and Calvo, R. 2023. What is a subliminal technique? An Ethical Perspective on AI Manipulation. *ETHICS-2023: IEEE International Symposium on Ethics in Engineering, Science, and Technology,* 1-10.

18. Chu, C., **Nyrup. R.**, Donato-Woodger. S., Leslie. K., Khan, S., Bernett, C. and Grenier, A. 2022. Examining the Technology-mediated Cycles of Injustice that Contribute to Digital Ageism. *PETRA '22: Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environment,* 545-551.

19. **Nyrup, R.** 2020. Three Uses of Analogy: A Philosophical View of the Archaeologist's Toolbox, *Interarchaeologia 6: Archaeology and Analogy*. Ed. M. Marila, M. Ahola, K. Mannermaa and M. Lavento. Helsinki: Department of Cultures, University of Helsinki, 12-31.

20. Whittlestone, J., **Nyrup, R.**, Alexandrova, A. and Cave, S. 2019. The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions, *AIES19: Proceedings of the 2019 AAAI/ACM Conference on Artificial Intelligence, Ethics and Society*, 195-200.

Book chapters

21. **Nyrup, R.** and Cibralic, B. forthcoming. Idealism, Realism, Pragmatism: Three Modes of Theorizing within Secular AI Ethics, *Research Handbook on Health, AI and the Law*. Eds. B. Solaiman and I.G. Cohen. Edward Elgar Publishing.

22. Keeling, G. and **Nyrup, R.** 2024. Explainable Machine Learning, Patient Autonomy, and Clinical Reasoning, *Oxford Handbook of Digital Ethics*, Ed. C. Véliz. Oxford UP.

23. **Nyrup, R.**, Chu, C. and Falco, E. 2023. Digital Ageism, Algorithmic Bias, and Feminist Critical Theory, *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Eds. J. Brown, S. Cave, K. E. Drage and K. McInnerney. Oxford UP.

24. **Nyrup, R.** 2021. Mortar and Pestle or Cooking Vessel? When Archaeology Makes Progress through Failed Analogies, *Explorations in Archaeology and Philosophy*, Synthese Library. Ed. A. Killin and S. Hermanson. Springer Nature.

Policy outreach and popular articles

25. **Nyrup, R.** Opening the black box: Transparency in AI is not enough. Institute of Arts and Ideas News, 13 October 2023.

26. Bermúdez, J.P., Calvo, R.A., **Nyrup, R.** and Deterding, S. The AI Act needs a practical definition of 'subliminal techniques'. Op-Ed in *Euractiv*, 1 Sep 2023.
27. **Nyrup, R.** and Scott-Fordsmand, H. 2022. Maskinlæring og uforklarlighedsproblemet. *Bibliotek for Læger* 4/2022.
28. Chu, C.H., Leslie, K., **Nyrup, R.**, Khan, S. 2022. Artificial intelligence can discriminate on the basis of race and gender, and also age, *The Conversation,* 18 January 2022.
29. **Nyrup, R.**, Whittlestone, J. and Cave, S. 2020. Why Value Judgements Should Not Be Automated. Written evidence submitted to the review into *Artificial Intelligence and Public Standards* by the UK Government's Committee on Standards in Public Life.
30. Whittlestone, J., **Nyrup, R.**, Alexandrova, A., Dihal, K. and Cave, S. 2019. *Ethical and Societal Implications of Algorithms, Data and Artificial Intelligence: A Roadmap for Research*. London: Nuffield Foundation.

Teaching materials

31. Rosenbaum, T., Møller, E., Topsøe, A. and **Nyrup, R.** 2011. *Kompendium i Retsfilosofi*, Frederiksberg: Kompendieforlaget. [Reading guide for Philosophy of Law course at the University of Copenhagen].

Book reviews and other

32. **Nyrup, R.** 2024. Review of Emmanuele Ratti and Thomas Stapleford: *Science, Technology and Virtues*. *Philosophy of Science* 91(1): 261-263.
33. Ganesh, M.I. and **Nyrup, R.** 2022. Cultural work in 'AI Ethics and Society' pedagogy: Some early reflections. *NeurIPS 2022*, Workshop on AI Cultures.
34. **Nyrup, R.** 2022. Mathematical string model, by Martin Schilling, German, late 19th Century - Wh.5175, *Materials for the History of Science: A Whipple Museum Festschrift to Celebrate the Directorship of Professor Liba Taub*. Ed. J. Nall, J. Hyslop and B. Jardine. Cambridge: Whipple Museum of the History of Science, 18-19.
35. **Nyrup, R.** 2018. Samuel Schindler: *Theoretical Virtues in Science*, *BJPS Review of Books*.

Recent talks

Dec 2023   General Principles vs. Procedural Values: Two Strategies of Socio-Ethical Regulation. *AI Risks and Regulation*, The Danish Society of Engineers (IDA).

Dec 2023   The Trouble with 'Trustworthy AI' (and Why We Need It), *CSS Colloquium*, Aarhus.

July 2023   Anthropomorphism without the hype? *Responsible AI Seminar Series*, Nokia Bell Labs, Cambridge.

Dec 2022   Explanatory Pragmatism and Philosophy for the Science of Explainable AI, *Departmental Seminar*, Department of STS, University College London.

Nov 2022   Rationality and pursuit in scientific theory choice, or Tom Kuhn and Tom Bayes meet Charlie Peirce, *Theoretical Virtues and Non-Empirical Guides in Scientific Theory Development*, Aarhus.

Oct 2022   What can Philosophy do for Responsible AI? *Responsible Engineering and Design Seminar*, Dyson School of Engineering, Imperial College London.

Oct 2022   Explanatory Pragmatism and Philosophy for the Science of Explainable AI, *Centre for Cognitive Sciences Research Seminar*, Sussex.

Jul 2022:   Explainable AI, Understanding and Non-Epistemic Values. *British Society for the Philosophy of Science* (BSPS), Exeter.

Apr 2022   Towards an Ethics of Explainable AI, *Ethical Engineering of Complex Systems: Robot Police, Judges, and Digital Beings*, Texas A&M.

Mar 2022   Towards an Ethics of Explainable AI, *Aarhus Conference on Algorithmic Decision Making*, Aarhus.

Feb 2022   Towards an Ethics of Explainable AI, *CHESS Seminar*, Department of Philosophy, Durham.

Nov 2021   Value Transparency and Value-Ladenness in Machine Learning. *Philosophy of Science Association Biennial Meeting* (PSA), Baltimore.

Grants and Awards

<u>Research grants</u>

2024-28   *Towards Responsible Explainable AI Technologies (TREAT)*, Sapere Aude Research Leader Grant, Independent Research Fund, Denmark. Role: PI.

2023-25   *Should I go to court? Improving Access to Justice Using Artificial Intelligence*, Nuffield Foundation. Role: Co-I. (PI: Felix Steffek, Cambridge).

2022-27   *Desirable Digitalisation: Rethinking AI for Just and Sustainable Futures*, Stiftung Mercator. Role: Researcher (PIs: Stephen Cave, U. Cambridge; Markus Gabriel, U. Bonn).

2020-22   *Ageism in A.I.: Exploring the Social and Ethical Implications of Age-Based Bias in Artificially Intelligent Systems*, Canadian Social Sciences and Humanities Research Council (SSHRC) Insight Development Grant. Role: Named collaborator (PI: Charlene Chu, U. Toronto).

2019-21   *Understanding Medical Black Boxes: A Philosophical Analysis of AI Explainability,* Humanities and Social Sciences Seed Award, Wellcome Trust. Role: PI.

<u>Studentships and other</u>

2016-17   Jacobsen Studentship, Royal Institute of Philosophy. PhD maintenance.

2015   Graduate Student Essay Prize (best paper), European Philosophy of Science Association Biennial Conference (EPSA2015).

2014-20:   Society for Applied Philosophy; British Society for the Philosophy of Science; British Society for History of Science; Aristotelian Society. Conference grants.

2014:   Augustinus Fonden; Oticon Fonden. External travel funding.

2013-16   Durham Doctoral Scholarship. Full-time PhD fees and maintenance.

2012-13:   AHRC Award, University of Cambridge. Full-time MPhil fees.

2011-13:   Julie Damms Studiefond. Studies abroad.

Outreach and Societal Impact

- Interview for article on the EU AI Act, *Akademikerbladet*, 28 Sep 2023.
- External consultant on industry-funded project (Digital Wellbeing Project, Imperial College London, PI: Rafael Calvo). Led work package on digital manipulation. 2022-23.
- Portrait interview in Danish newspaper, *Weekendavisen*, 20 Nov 2022.
- Interviewed on AI risks for article in Danish newspaper, *Politiken*, 6 Aug 2022.
- Interview on Values in Artificial Intelligence, *Philosophy Exchange Podcast*, 3 March 2021.
- Member of Health and Social Care Advisory Panel, Centre for Data Ethics and Innovation, UK Government, 2019.
- Provided written and oral evidence to the Committee on Standards in Public Life's review into *Artificial Intelligence and Public Standards*, UK Government 2019.
- Advised and reviewed on report algorithmic bias, Centre for Data Ethics and Innovation, UK Government. 2019.
- Advised on the ethics of autonomous vehicles for the exhibition "Driverless: who is in control", London Science Museum, 2018-2019.
- Chaired Panel on "Population Health Data and AI: The Ethics", *UK e-Health Week* (Digital healthcare event for industry, policy makers and service providers), 16 May 2018.
- Interview on ethical issues of AI for Danish newspaper, *Kristeligt Dagblad*, 27 Dec 2017.
- Appeared on BBC Radio 4's *Moral Maze* on "the Morality of AI", 22 Nov 2017

(see also 'Policy outreach and popular articles' above)

Teaching

Lecturing and postgraduate seminars
- *Philosophy of Science and Ethics for Computer Science, IT-Product Development and Data Science*. Aarhus, 2024.
- *Explainable AI*, MSc in Science Studies, Aarhus, 2024.
- *AI in Healthcare*, MPhil in Health, Medicine and Society, Department of HPS, Cambridge. 2021-23.
- *Machine Ethics*, MSt in AI Ethics and Society, Institute for Continuing Education, Cambridge. 2022.
- *Philosophical Reflections on an AI Future*, summer school with the Swiss Study Foundation. 2019.
- *Philosophy of Science*, Department of Philosophy, Durham. Guest lectures/seminars. 2015, 2016.
- *Philosophy of Science for Biologists*, Department of Biology, Copenhagen. Guest lecture. 2012.

Dissertation and Postgraduate supervising
- 4 dissertations, MSt in AI Ethics and Society, Institute of Continuing Education, Cambridge.
- 13 MPhil essays and 4 dissertations. Department of HPS and Faculty of Philosophy, Cambridge.
- 4 undergraduate dissertations (3 as primary supervisor, 1 as second). HPS and Philosophy, Cambridge.
- 5 Summer Online Research Programme dissertations (UG level), Pembroke College, Cambridge.
- 2 Advanced Diploma dissertations (UG level). Institute of Continuing Education, Cambridge.

Other postgraduate teaching
- Academic Adviser, MSt in AI Ethics and Society, Institute of Continuing Education, Cambridge. 2021-23.
- Work in progress group for MPhil students, Department of HPS, Cambridge. 2021.

Undergraduate seminars/tutorials
- *Ethics and Political Philosophy*, Faculty of Philosophy, Cambridge. 2018.
- *History and Philosophy of Science IB*, Department of HPS, Cambridge. 2017, 2018.
- *Philosophy of Science.* Faculty of Philosophy, Cambridge. 2017, 2018.
- *Politics and Ethics of Medicine*, Department of HPS, Cambridge. 2017.
- *Moral Theory*, Department of Philosophy, Durham. 2016-17.
- *Introduction to HPS*, Department of Philosophy, Durham. 2015-16.
- *Knowledge and Reality,* Department of Philosophy, Durham. 2014-15.
- *Philosophy of Science for Biologists*, Department of Biology, Copenhagen. 2012.
- *Philosophy of Science and Ethics for Physicists*, Niels Bohr Institute, Copenhagen. 2011

Professional Service

Administration and institutional service
- Postgraduate Admissions Officer, MPhil in Ethics of AI, Data and Algorithms, Leverhulme Centre for the Future of Intelligence, University of Cambridge. 2022-23.
- Led the successful development of a new MPhil in Ethics of AI, Data and Algorithms, Leverhulme Centre for the Future of Intelligence, University of Cambridge. 2021-22.
- Researcher Representative (2020-23) and Deputy Researcher Representative (2019-20), Leverhulme Centre for the Future of Intelligence, Cambridge.
- Wellbeing, Inclusion, Diversity and Equality (WIDE) Group, Leverhulme Centre for the Future of Intelligence, Cambridge. Secretary 2019-20; Ordinary member 2021-22; Chair 2022-23.
- Centre Representative, School of Arts and Humanities Equality, Diversity and Inclusion Forum, Cambridge, 2022.
- Examiner, MPhil/Part III dissertations and essays, Department of HPS, Cambridge, 2017-2023.
- Director of Studies in History and Philosophy of Science, Churchill College. 2017-20; 2022-23.
- Member of the Board and the Degree Committee, Department of HPS, Cambridge. Jan 2018-Dec 2019.

- Managed and supervised 2 postdocs on the Digital Wellbeing project (Imperial College, 2022-2023).
- Managed and supervised 2 part-time research assistants (both PhD-students) on my Medical Black Boxes Project (Cambridge, 2019-2021).
- Hiring committee member for 7 positions (2 postdocs, 3 research assistants, 2 administrators). Shortlisting and interviews. Cambridge, Imperial College, Aarhus University.

Organising

July 2022:	Symposium: Assisted Discoveries, Understanding and the Aims of Science, *British Society for the Philosophy of Science*, Exeter. (w/ M. Ivanova and M. Stuart)
Nov 2021:	Symposium: Philosophy of Science Meets AI Ethics, *Philosophy of Science Association*, Baltimore. (w/ J. Biddle, K. Vredenburgh and E. Sullivan).
May 2021:	Understanding and Explaining in Healthcare, Cambridge. Part of the *Issues in Explainable AI* series. (w/ G. Keeling & E. Falco)
Jun 2019	Representing and Intervening: AI Ethics Meets Philosophy of Science, Cambridge. (w/ S. John)
Sep 2018	Values in Science: New Perspectives and Directions, Cambridge. (w/ A. Alexandrova and S. John)
Sep 2018	Philosophy and the Ethics of Algorithmic Decision Making, Cambridge (w/ A. Alexandrova and S. John).
Dec 2017	Games, Values and AI, Cambridge. (w/ H. Shevlin)
2017-18	Philosophy of the Historical Sciences Reading Group, Cambridge. (w/ A. Ion and A. Currie)
Sep 2017	Values and AI in Healthcare and Medicine, Cambridge. (w/ A. Alexandrova and S. John)
Oct 2014	Philosophy and Public Policy, Durham. (w/ A. de Bruyckere)

Peer reviewing

- *Philosophy:* Biology and Philosophy; British Journal for the Philosophy of Science; British Society for Philosophy of Science 2018 Annual Conference (programme committee); Cambridge Quarterly of Healthcare Ethics; Canadian Journal of Philosophy; Erkenntnis; Clinical Ethics; Journal of Evaluation in Clinical Practice; Minds and Machines; Philosophy & Technology; Philosophy of Science; Oxford University Press (chapter in edited volume); Oxford Bibliographies; Studies in History and Philosophy of Science; Synthese; Transactions of the C.S. Peirce Society.
- *Interdisciplinary AI Ethics:* ACM FAccT 2022 (programme committee); ACM/AAAI Conference on AI Ethics and Society 2020 (full paper); ACM Computing Surveys; Human Brain Mapping; IEEE Transactions on Affective Computing; IEEE Transactions on Technology and Society; Journal of Cyber Policy; KI2019—German Conference on Artificial Intelligence (full paper); Proceedings of the IEEE.

Miscellaneous

- Seminar Blog contributor, Centre for Humanities Engaging Science and Society, Durham (https://www.dur.ac.uk/chess/seminarblog/), 2014-2016.
- Founding co-editor, *Tanken* (philosophy student magazine). University of Copenhagen, (https://tankenonline.dk). 2009-12.