David Watson

20a Mildmay Grove South • London, N1 4RL
Phone: +44 (0)7478 730102 • Email: david.s.watson11@gmail.com

Education

University of Oxford **October 2017—January 2021**
- DPhil in Information, Communication, and the Social Sciences **Oxford, UK**
- Developed formal frameworks for explainable artificial intelligence
- Implemented flexible algorithms for global and local model explanations

The Alan Turing Institute **October 2018—September 2019**
- Doctoral Enrichment Student **London, UK**
- Helped draft ICO's *Project ExplAIn Guidance*
- Led weekly reading group on statistical learning theory

University of Oxford **October 2014—August 2015**
- MSc in Social Science of the Internet **Oxford, UK**
- Studied big data analytics, internet economics, and information visualization
- Wrote a thesis on crowdsourcing in the natural sciences

Dartmouth College **September 2007—June 2011**
- BA in Philosophy, High Honors **Hanover, NH**
- Studied philosophy of science, linguistics, and literature
- Wrote a thesis on the metaphysics of quantum cosmology

Professional Experience

Postdoctoral Research Fellow, University College London **January 2021—**
- Conducting original research on causality and machine learning **London, UK**
- Co-supervising doctoral candidates at the Centre for Artificial Intelligence
- Lecturing on graphical models in the Department of Statistical Science

Data Scientist, Queen Mary University of London **November 2015—December 2020**
- Conducted exploratory and inferential analytics for bioinformatics projects **London, UK**
- Developed unsupervised learning algorithms for genomic data integration
- Created visualization software for gene expression studies

Research Assistant, Int'l Assoc for Computing & Philosophy **March 2019—April 2020**
- Developed formal models of explanation in artificial intelligence **Oxford, UK**
- Conducted literature review on philosophical foundations of machine learning
- Drafted articles on the epistemology of data science

Teaching Assistant, University of Oxford **October 2018—December 2020**
- Lectured on the philosophy and ethics of information **Oxford, UK**
- Led seminars on the foundations of social data science
- Developed curricula for future MSc cohorts

Freelance Contributor, *The Economist* **March 2016—October 2019**
- Wrote articles for the Graphic Detail section and Game Theory blog **London, UK**
- Built simulations to estimate the probability of global events
- Collaborated with editorial staff to research and develop new stories

Assistant Editor, HarperCollins Publishers **December 2011—September 2014**

- Read and reviewed manuscripts for publication **New York, NY**
- Launched and managed e-book classics program
- Appointed Editorial Director of the National Poetry Series

SELECT PUBLICATIONS

Watson, D. & Wright, M. (2021). Testing conditional independence in supervised learning algorithms. *Machine Learning*.

Watson, D., Gultchin, L., Taly, A., & Floridi, L. (2021). Local explanations via necessity and sufficiency: Unifying theory and practice. In *International Conference on Uncertainty in Artificial Intelligence*.

Gultchin, L., Watson, D., Kusner, M., & Silva, R. (2021). Operationalizing complex causes: A pragmatic view of mediation. In *International Conference on Machine Learning*. Vienna, Austria.

Kinney, D. & Watson, D. (2020). Causal feature learning for utility-maximizing agents. In *International Conference on Probabilistic Graphical Models* (pp. 257–268). Skørping, Denmark.

Nicholls, H.L., John, C.R., Watson, D., Munroe, P.B., Barnes, M.R., & Cabrera, C.P. (2020). Reaching the end-game for GWAS: Machine learning approaches for the prioritization of complex disease loci. *Frontiers in Genetics*, *11*, 350.

Watson, D. & Floridi, L. (2020). The explanation game: A formal framework for interpretable machine learning. *Synthese*.

John, C.R., Watson, D., Russ, D., Goldmann, K., Ehrenstein, M., Pitzalis, C., … Barnes, M. (2020). M3C: Monte Carlo reference-based consensus clustering. *Scientific Reports*, *10*(1), 1816.

Watson, D. (2019). The rhetoric and reality of anthropomorphism in artificial intelligence. *Minds & Machines*, *29*(3), 417-440.

John, C.R., Watson, D., Barnes, M.R., Pitzalis, C., & Lewis, M. (2019). Spectrum: Fast density-aware spectral clustering for single and multi-omic data. *Bioinformatics*, *36*(4), 1159–1166.

Watson, D. (2019). The price of discovery: A model of scientific research markets. In Öhman, C. & Watson, D. (Eds.), *The 2018 Yearbook of the Digital Ethics Lab* (pp. 51–63). Heidelberg: Springer.

Öhman, C. & Watson, D. (Eds.) (2019). *The 2018 Yearbook of the Digital Ethics Lab*. Heidelberg: Springer.

Öhman, C. & Watson, D. (2019). Are the dead taking over Facebook? A big data approach to the future of death online. *Big Data & Society*, *6*(1), 1-13.

Watson, D., Krutzinna, J., Bruce, I.N., Griffiths, C.E.M., McInnes, I.B., Barnes, M.R., & Floridi, L. (2019). Clinical applications of machine learning algorithms: Beyond the black box. *BMJ*, *364*.

O'Toole, S.M., Watson, D., Novoselova, T.V., Romano, L.E.L., King, P., Bradshaw, T.Y., … Chapple, J.P. (2019). Oncometabolite induced primary cilia loss in pheochromocytoma. *Endocrine-Related Cancer*, *26*(1), 165-180.

Watson, D. & Floridi, L. (2018). Crowdsourced science: Sociotechnical epistemology in the e-research paradigm. *Synthese*, *195*(2), 741–764.

Foulkes, A.C., Watson, D., Carr, D.F., Kenny, J.G., Slidel, T., Parslew, R., … Barnes, M.R. (2018). A framework for multi-omic prediction of treatment response to biologic therapy for psoriasis. *Journal of Investigative Dermatology*, *139*(1), 100–107.

Cabrera, C.P., Manson, J., Shepherd, J.M., Torrance, H.D., Watson, D., Longhi, M.P., … Brohi, K. (2017). Signatures of inflammation and impending multiple organ dysfunction in the hyperacute phase of trauma: A prospective cohort study. *PLOS Medicine*, *14*(7), e1002352.

Foulkes, A.C., Watson, D., Griffiths, C.E.M., Warren, R.B., Huber, W., & Barnes, M.R. (2017). Research techniques made simple: Bioinformatics for genome-scale biology. *Journal of Investigative Dermatology*, *137*(9), e163–e168.

Preprints

Watson, D. (2021). Rational Shapley values. *arXiv* preprint, 2106.10191.

Watson, D. (2020). Conceptual challenges for interpretable machine learning. SSRN preprint, 3668444.

Software

Watson, D. & Tansey, W. (2021). smoothFDR: An empirical Bayes method for exploiting spatial structure in large multiple-testing problems. URL: https://github.com/dswatson/smoothFDR.

Watson, D. (2021). bioplotr: Pretty, simple, optionally interactive plots for bioinformatics analysis pipelines. URL: https://github.com/dswatson/bioplotr.

Watson, D. & Wright, M. (2021). cpi: Testing conditional independence in supervised learning algorithms. URL: https://github.com/dswatson/cpi.

John, C.R. & Watson, D. (2021). M3C: Monte Carlo reference-based consensus clustering. URL: https://bioconductor.org/packages/release/bioc/html/M3C.html.

John, C.R. & Watson, D. (2021). Spectrum: Fast adaptive spectral clustering for single and multi-view data. URL: https://cran.r-project.org/package=Spectrum.

Select Presentations

'Interpretable machine learning for genomics: Opportunities, challenges, and future directions.' Understanding and Explaining in Healthcare. Cambridge, May 2021.

'No explanation without inference: Algorithmic opacity and severe testing.' AISB Symposium on Opacity in Machine Learning. London, April 2021.

'Necessary and sufficient factors for contrastive local explanation.' The Digital Ethics Lab, University of Oxford, January 2021.

'Pragmatic causal feature learning.' International Conference on Probabilistic Graphical Models. Aalborg, September 2020.

'Conceptual challenges for interpretable machine learning.' ACM Conference on Fairness, Accountability, and Transparency in Machine Learning, Doctoral Consortium. Barcelona, January 2020.

'Machine learning for predicting clinical outcomes.' PSORT Showcase. Royal College of Physicians, London, November 2019.

'Information ethics: Theories, problems, strategies.' Learn, develop & design: Ethics principles through cross-disciplinary collaboration. Royal College of Art, London, September 2019.

'The explanation game: A formal framework for interpretable machine learning.' 12th Annual MuST Conference on Statistical Reasoning and Scientific Error. Ludwig Maximilian University, Munich, July 2019.

'Interpretable machine learning for clinical medicine.' Mining Science Data for Medicine Workshop. University of Manchester, April 2019.

'The rhetoric and reality of anthropomorphism in artificial intelligence.' The Digital Ethics Lab, University of Oxford, January 2019.

'Attention economies and the ethics of design.' London Doctoral Design Centre, Royal College of Art, London, November 2018.

'High-dimensional model explanations with applications to genomics.' The Alan Turing Institute, London, April 2018.

'Formal frameworks for interpretable machine learning.' The Digital Ethics Lab, University of Oxford, November 2017.

‘The EAGLE has landed: Real-time win probabilities in men’s major golf tournaments.’ MIT Sloan Sports Analytics Conference. Hynes Convention Center, Boston, February 2017.

‘Omics primer for clinicians: An introduction to high-dimensional statistics.’ British Association of Dermatologists Workshop. University of Manchester, December 2016.

‘Measuring the epistemological and social impact of citizen science.’ St. Anne’s College, University of Oxford, December 2016.

‘Modelling biologic response: Clinical and statistical considerations.’ Stratified Medicine Workshop. Francis Crick Institute, London, October 2016.

‘Feature selection in high-dimensional classification problems.’ CSAMA Conference on Statistical Data Analysis for Genome Scale Biology. University of Padua, Brixen, July 2016.