Matthew Evans

matthew@ml-evs.science
ml-evs.science
ml-evs

decentralized data management • open science & software
materials discovery • ab initio calculations

RESEARCH INTERESTS

My background in computational materials science has left me with an overarching interest in the application of machine learning, open source software & infrastructure, and data management practices to accelerate and enhance scientific workflows for discovery in the chemical and materials sciences.

EDUCATION

2016–2023 **PhD Physics**, *(submitted July 2023)*, Theory of Condensed Matter Group, University of Cambridge

2015–2016 **MPhil Scientific Computing**, University of Cambridge, *Pass with distinction*

2011–2015 **MPhys Physics with Theoretical Physics**, University of Manchester, *First Class (Hons)*

SELECTED EXPERIENCE

2025– **Leverhulme Early Career Research Fellow**

University of Cambridge, with Prof Clare Grey FRS

2024– **Scientific Software Consultant** and **Director**

datalab industries ltd.

- Supporting the open source development of OPTIMADE and *datalab* via consultancy services.
- Customisation and deployment of *datalab* for industrial R&D and academic labs.

2020–2025 **Research Assistant** (2020–2022) then **BEWARE Research Fellow** (2022–2025)

IMCN, Université catholique de Louvain and Matgenix SRL., with Prof Gian-Marco Rignanese

- Co-creator and architect of *datalab*, open source data management software for sample tracking and characterisation, lab management, and machine learning, deployed at several labs internationally.
- High-throughput machine-learning accelerated workflows for materials discovery and design.
- Leading development of the OPTIMADE API specification and associated software.

2016–2020 **PhD student: Crystal structure prediction for next-generation energy storage**

University of Cambridge, with Dr Andrew Morris

- Computional materials discovery for conversion anodes for Li, Na and K-ion batteries.
- Author of two open-source Python packages: database approaches for high-throughput calculations and materials design with `matador` and crystal structure prediction with `ilustrado`.

COMPUTING

Languages **Python**, Javascript, Vue.js, Fortran, C++

Practices **Test-driven development, CI/CD, Cloud Automation, HPC**

Tools **git**, **vim**, **Docker**, **Ansible**, **Terraform**

Expertise Web APIs, databases, machine learning, high-throughput workflows

OTHER EXPERIENCE

2021–2025 **Visiting researcher**: Data management platforms for materials chemistry research
Department of Chemistry, University of Cambridge with Prof Clare Grey

2022 **Postdoctoral Researcher** Recommender systems for crystal structure search
Cambridge Crystallographic Data Centre, Cambridge, UK

2019 **Visiting Researcher**: Machine learning for materials discovery
Aalto University, Finland, with Profs Adam Foster & Patrick Rinke

2019 **Scientific Software Developer (Intern):** Multi-objective optimisation
Enthought Inc., Cambridge

2014, 2015 **UG research**: Interactions of quantised vortices in superfluid helium
University of Manchester, with Dr Paul Walmsley & Prof Andrei Golov

2013 **UG research**: Hard sphere packing of nanotube-encapsulated fullerenes
University of Nottingham, with Dr Ho-Kei Chan & Prof Elena Besley

SELECTED (AWARDS + HONOURS)

2025 Leverhulme Early Career Research Fellowship with additional support from Isaac Newton Trust.

2022 BEWARE2 Fellowship from the Wallonia-Brussels Federation (approx. €300,000).

2021 PI for "Interoperable data management for fundamental battery research", BIG-MAP External Stakeholder Initiative (approx. €150,000).

SELECTED (TEACHING + SERVICE)

2018– Reviewed for *Digital Discovery* (x7), *JOSS* (x6), *J. Phys.: Cond. Mat.* (x4), *Mach. Learn.: Sci. Technol.* (x2) (IOP Outstanding Reviewer 2024), *npj. Comp. Mater.* (x1), *Sci Data* (x1), *Sci. Rep.* (x1), *Sci. Technol. Adv. Mater.: Methods* (x1).

2022–2024 Proposed and co-lead a MaRDA working group on metadata extractors for materials science.

2022–2024 Initiator and organiser of the CECAM Workshop series *Machine-actionable Data Interoperability for Chemical Sciences* (MADICES, February 2022, April 2024 and upcoming in October 2025)

2019-2021 Demonstrator: Part II Computational Physics, 3x Part IB Intro to Computing, Cavendish Laboratory

2016–2018 Supervisor: 2x Part IB Electromagnetism, Dynamics and Thermodynamics, Selwyn College

2012–2015 Tutor: GCSE Maths & Key Stage 2 Programming for The Tutor Trust, Manchester

SELECTED RECENT PRESENTATIONS

2025 Invited talk: *Decentralized materials research data management, curation and dissemination for accelerated discovery*, Machine Learning for Spectroscopy, Vrije Universiteit Brussel, Belgium.

4x Invited seminars: *Decentralized materials research data management, curation and dissemination for accelerated discovery*. Materials Modelling Discussion Group, University of Birmingham; PSDI Polymer Data Workshop, Loughborough University; Computational Chemistry Seminar Series, University of Warwick; AIChemy Seminar Series, Imperial College London

2024 Invited talk: *Decentralized materials research data management, curation and dissemination for accelerated discovery*, Symposium: Democratizing AI in Materials Science, MRS Fall, Boston, USA

Invited talk: *Federated, interoperable databases for accelerated materials discovery and design*, CECAM Flagship Workshop on MLIPs and Accessible Databases, Grenoble, France.

PUBLICATIONS

Underline indicates (joint) first authorship (reordered where appropriate). Full list with OA links available online (https://ml-evs.science/papers, ORCiD, Google Scholar).

Preprints

15. Zimmermann, Y., Bazgir, A., *et al. Reflections from the 2024 Large Language Model (LLM) Hackathon for Applications in Materials Science and Chemistry* (submitted). 2024. 10.48550/arXiv.2411.15221.

Peer-reviewed articles

20. Zimmermann, Y., Bazgir, A., *et al.* 32 Examples of LLM Applications in Materials Science and Chemistry: Towards Automation, Assistants, Agents, and Accelerated Scientific Discovery. *Mach. Learn.: Sci. Technol.* (accepted), (2025). 10.48550/arXiv.2505.03049.

19. Trinquet, V., **Evans, M. L.** & Rignanese, G.-M. Accelerating the discovery of high-performance nonlinear optical materials using active learning and high-throughput screening. *J. Mat. Chem. C,* (2025). 10.1039/d5tc01335f.

18. **Evans, M. L.**, Rignanese, G.-M., Elbert, D. & Kraus, P. Metadata, automation, and registries for extractor interoperability in the chemical and materials sciences. *MRS Bulletin,* (2025). 10.1557/s43577-025-00925-8.

17. Mroz, A. M., **Evans, M. L.**, *et al.* Cross-disciplinary perspectives on the potential for artificial intelligence across chemistry. *Chemical Society Reviews,* (2025). 10.1039/D5CS00146C.

16. **Evans, M. L.**, Trinquet, V., *et al.* Optical materials discovery and design with federated databases and machine learning. *Faraday Discussions* **256,** 459–482, (2025). 10.1039/D4FD00092G.

14. **Evans, M. L.**, Bergsma, J., Merkys, A., Andersen, C. W., *et al.* Developments and applications of the OPTIMADE API for materials data exchange and discovery. *Digital Discovery* **3,** (2024). 10.1039/D4DD00039K.

13. Rosen, A. S., Gallant, M., *et al.* Jobflow: Computational Workflows Made Simple. *Journal of Open Source Software* **9,** 5995, (2024). 10.21105/joss.05995. (2024).

12. Wang, Z., Gong, Y., **Evans, M. L.**, *et al.* Machine learning-accelerated discovery of A_2BC_2 ternary electrides with diverse anionic electron densities. *J. Amer. Chem. Soc.* **145,** 26412–26424, (2023). 10.1021/jacs.3c10538.

11. Lertkiattrakul, M., **Evans, M. L.** & Cliffe, M. J. PASCal Python: A Principal Axis Strain Calculator. *Journal of Open Source Software* **8,** 5556, (2023). 10.21105/joss.05556.

10. Jablonka, K. M., Ai, Q., *et al.* 14 Examples of How LLMs Can Transform Materials Science and Chemistry: A Reflection on a Large Language Model Hackathon. *Digital Discovery,* (2023). 10/gswbnx.

9. Ells, A. W., **Evans, M. L.**, Groh, M., Morris, A. J. & Marbella, L. E. Phase transformations and phase segregation during potassiation of Sn_xP_y anodes. *Chemistry of Materials,* (2022). 10/h69d.

8. **Evans, M. L.**, Andersen, C. W., *et al.* optimade-python-tools: a Python library for serving and consuming materials data via OPTIMADE APIs. *Journal of Open Source Software* **6,** 3458, (2021). 10/gn3w9f.

7. **Evans, M. L.**, Andersen, C. W., *et al.* OPTIMADE, an API for exchanging materials data. *Scientific Data* **8,** 217, (2021). 10/gmnrxj.

6. Breuck, P.-P. D., **Evans, M. L.** & Rignanese, G.-M. Robust model benchmarking and bias-imbalance in data-driven materials science: a case study on MODNet. *J. Phys.: Cond. Mat.* **33,** 404002, (2021). 10/gpw93d.

5. **Evans, M. L.** & Morris, A. J. matador: a Python library for analysing, curating and performing high-throughput density-functional theory calculations. *Journal of Open Source Software* **5,** 2563, (2020). 10/gmf4mv.

4. Harper, A. F., **Evans, M. L.** & Morris, A. J. Computational Investigation of Copper Phosphides as Conversion Anodes for Lithium-Ion Batteries. *Chemistry of Materials,* (2020). 10/gg5sx3.

3. Harper, A. F., **Evans, M. L.**, *et al.* Ab initio Structure Prediction Methods for Battery Materials. *Johnson Matthey Technology Review* **64,** 103–118, (2020). 10/ggrmgf.

2. Marbella, L. E., **Evans, M. L.**, *et al.* Sodiation and Desodiation via Helical Phosphorus Intermediates in High-Capacity Anodes for Sodium-Ion Batteries. *J. Amer. Chem. Soc.* **140,** 7994–8004, (2018). 10/gdq6h4.

1. Zhu, T., **Evans, M. L.**, Brown, R. A., Walmsley, P. M. & Golov, A. I. Interactions between unidirectional quantized vortex rings. *Physical Review Fluids* **1,** 044502, (2016). 10/gf2529.