EMOS: The Unified AI Platform for Electronic Materials Discovery

Atish Dixit [1] **Alexandros Keros** [1] **Ben D. Rowlinson** [1] **Shashank Mishra** [1] **Jacqui Cole** [2]
Subramanian Ramamoorthy [3] **Themis Prodromakis** [4]

[1] *APRIL AI Hub, University of Edinburgh, Edinburgh EH9 3BF, UK* [2] *Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK* [3] *Institute of Perception, Action and Behaviour, School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, UK* [4] *Centre for Electronic Frontiers, School of Engineering, University of Edinburgh, Edinburgh EH9 3BF, UK*. Correspondence to: Atish Dixit adixit@ed.ac.uk.

1. Introduction

AI-based materials discovery is gaining significant traction in electronic materials research, driven by three main families of computational tools: i) curated databases (e.g., Materials Project [1], The Alexandria Library [2], MatHub-2d [3]) provide structural and property data, ii) generative models (e.g., Mattergen [4], iMatGen [5], CDVAE [6]) propose novel material candidates, and iii) property predictors (e.g., MatterSim [7], Orb [8], MACE [9], ALIGNN [10]) guide screening and assessment. Designing electronic devices requires integrating multiple computational models to satisfy multi-objective targets including metallicity, band gap, thermal conductivity, and carrier mobility. EMOS serves as a foundational platform for such materials discovery and device design.

The current landscape remains fragmented: databases differ in schema and coverage, models target diverse descriptors and properties, and materials discovery workflows lack standard interfaces across the pipeline. This forces researchers to develop brittle glue code, leading to duplicated effort and poor reproducibility. Consequently, two critical challenges persist: integrating heterogeneous models and databases, and building customizable pipelines tailored to electronic materials discovery and device design.

Open platforms with clear governance structures are needed to lower barriers to integration, standardize metadata, and enable reuse. While prior initiatives have made progress in specific areas, Table 1 shows that no existing framework provides the comprehensive integration necessary for electronic materials workflows.

EMOS addresses this gap by unifying data, generators, and predictors into composable *information units* within a consistent interface. Researchers can assemble *features*, or custom workflows, for electronic device design without bespoke adapters. New units and features can be contributed and reused, while shared logging and provenance support benchmarking and reproducibility. By prioritizing interoperability and modularity, EMOS transforms fragmented assets into deployable, end-to-end AI workflows for electronic materials discovery.

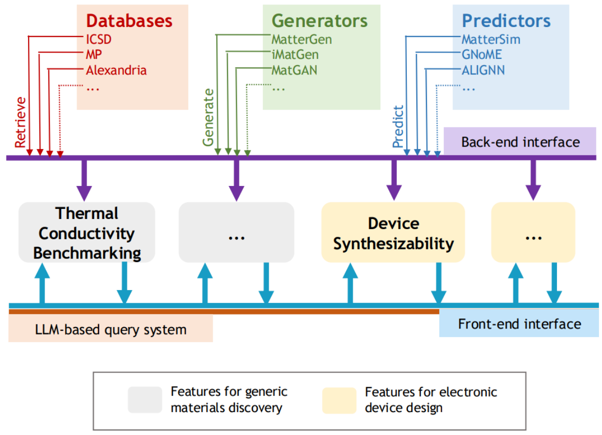

Fig. 1: EMOS framework architecture with core abstractions: information units, features, and their interconnections

2. EMOS Framework

EMOS is an open-source, forkable framework with documentation and examples for onboarding, enabling feedback and refinement via issues and pull requests, while an interactive web application provides access without local installation.

The framework is organized around three core abstractions: information units, features, and their interconnections (Figure 1). Information units encapsulate materials databases, generative models, and property predictors behind standardized APIs, while features compose these units into end-to-end workflows for electronic materials and device design. By explicitly representing connections as a graph, EMOS

Table 1: Comprehensive integration in EMOS versus existing materials discovery platforms

Platform	Database Aggregation	Generators Aggregation	Predictors Aggregation	Benchmarking Capability	Device Design Customization
JARVIS [11]	✓	✗	✗	✗	✗
OPTIMADE [12]	✓	✗	✗	✗	✗
MatBench [13]	✗	✗	✗	✓	✗
CHIPS-FF [14]	✗	✗	✓	✓	✗
EMOS	✓	✓	✓	✓	✓

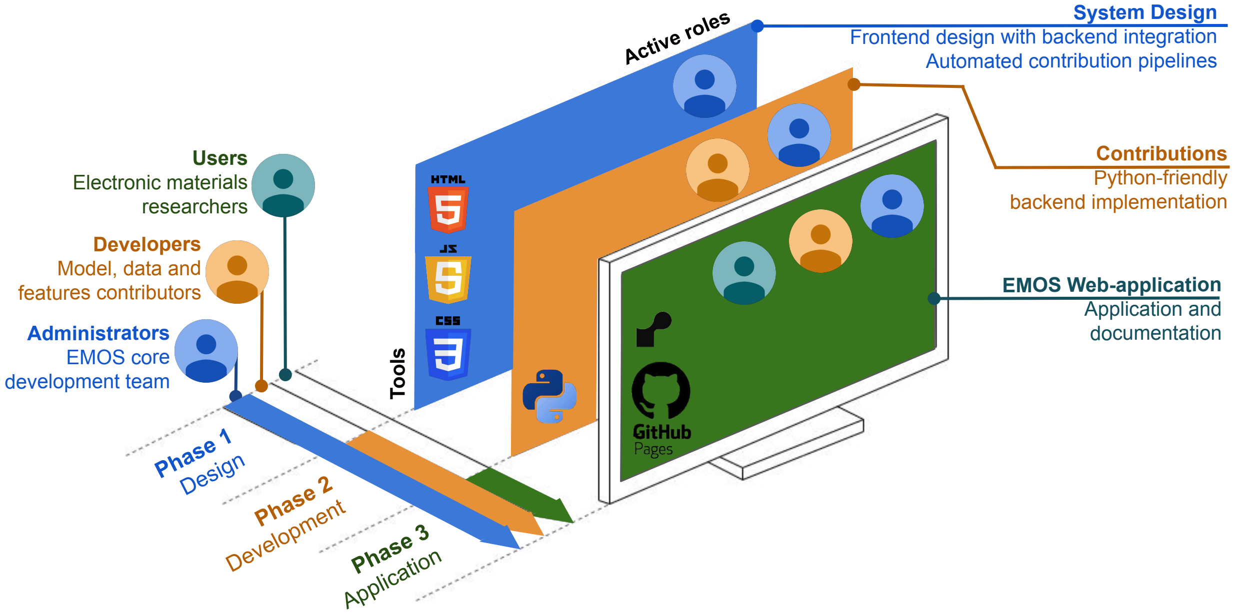

Fig. 2: System design showing contributor roles on left and deployment phases as layered panels leading to the EMOS web-application on right

enables generators and predictors to be interchanged without reimplementing downstream components, enabling systematic model comparison and component reuse.

EMOS distinguishes contributor roles, deployment phases, and the front-end/back-end split (Figure 2), allowing the platform to evolve while maintaining coherence and quality.

The front end (HTML, CSS, JavaScript) provides a consistent interface, while the Python back end exposes APIs for data access and orchestration, with the web client hosted on GitHub Pages and compute tasks delegated to Render.

Contributions remain lightweight: developers introduce new units or features via Python APIs without coupling backend to frontend. Admin review provides quality control, while community feedback guides alignment with shared standards, supporting sustained, community-driven development of EMOS.

3. Result: an Application Scenario

Figure 3 illustrates an application of EMOS framework for MOSFET device design, demonstrating feature assembly into complete pipelines. The workflow chains database resources through generative models and predictors with domain-specific filters, progressively refining candidates according to MOSFET specifications, demonstrating how EMOS enables tailored pipelines for specific device classes and performance targets.

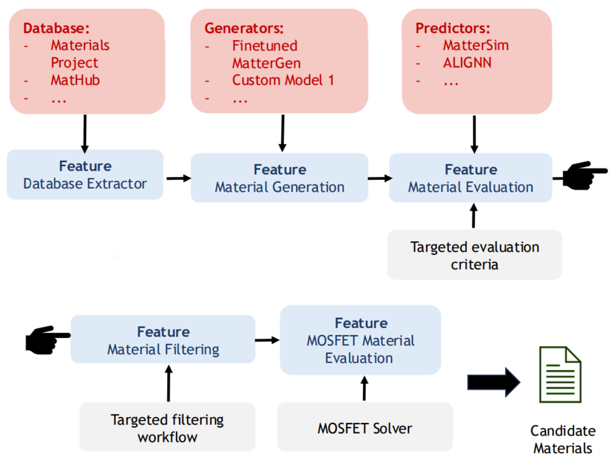

Fig. 3: An application scenario of AI-based materials discovery for MOSFET device design

Code Availability

The source code is available at `https://github.com/aprilaihub/EMOS`. An interactive web interface is publicly accessible at `https://aprilaihub.github.io/EMOS/` and through the APRIL website (`https://www.april.ac.uk/`, under *Research → EMOS App*).

Acknowledgments

We acknowledge support of the UKRI AI Hub programme for the APRIL AI Hub [Grant number EP/Y029763/1].

References

[1] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.

[2] Mohammad M Ghahremanpour, Paul J Van Maaren, and David Van Der Spoel. The alexandria library, a quantum-chemical database of molecular properties for force field development. *Scientific data*, 5(1):1–10, 2018.

[3] Mingjia Yao, Jialin Ji, Xin Li, Zhenyu Zhu, Jun-Yi Ge, David J Singh, Jinyang Xi, Jiong Yang, and Wenqing Zhang. Mathub-2d: A database for transport in 2d materials and a demonstration of high-throughput computational screening for high-mobility 2d semiconducting materials. *Science China Materials*, 66(7):2768–2776, 2023.

[4] Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.

[5] Juhwan Noh, Jaehoon Kim, Helge S Stein, Benjamin Sanchez-Lengeling, John M Gregoire, Alan Aspuru-Guzik, and Yousung Jung. Inverse design of solid-state materials via a continuous representation. *Matter*, 1(5):1370–1384, 2019.

[6] Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197*, 2021.

[7] Han Yang, Chenxi Hu, Yichi Zhou, Xixian Liu, Yu Shi, Jielan Li, Guanzhi Li, Zekun Chen, Shuizhou Chen, Claudio Zeni, et al. Mattersim: A deep learning atomistic model across elements, temperatures and pressures. *arXiv preprint arXiv:2405.04967*, 2024.

[8] Mark Neumann, James Gin, Benjamin Rhodes, Steven Bennett, Zhiyi Li, Hitarth Choubisa, Arthur Hussey, and Jonathan Godwin. Orb: A fast, scalable neural network potential. *arXiv preprint arXiv:2410.22570*, 2024.

[9] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.

[10] Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, 2021.

[11] Kamal Choudhary, Kevin F Garrity, Andrew CE Reid, Brian DeCost, Adam J Biacchi, Angela R Hight Walker, Zachary Trautt, Jason Hattrick-Simpers, A Gilad Kusne, Andrea Centrone, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj computational materials*, 6(1):173, 2020.

[12] Casper W Andersen, Rickard Armiento, Evgeny Blokhin, Gareth J Conduit, Shyam Dwaraknath, Matthew L Evans, Ádám Fekete, Abhijith Gopakumar, Saulius Gražulis, Andrius Merkys, et al. Optimade, an api for exchanging materials data. *Scientific data*, 8(1):217, 2021.

[13] Janosh Riebesell, Rhys EA Goodall, Philipp Benner, Yuan Chiang, Bowen Deng, Alpha A Lee, Anubhav Jain, and Kristin A Persson. Matbench discovery–a framework to evaluate machine learning crystal stability predictions. *arXiv preprint arXiv:2308.14920*, 2023.

[14] Daniel Wines and Kamal Choudhary. Chipsff: Evaluating universal machine learning force fields for material properties. *ACS Materials Letters*, 7(6):2105–2114, 2025.