AI4Materials: A Manifesto for AI-Driven Scientific Discovery in Materials Science*

Eros Radicchi[1,2], Usman Syed[1], Federico Cunico[1], Michele Cassetta[3,1], Uzair Khan[1], Paolo Marone[4], Alberto Scarsini[5], Filiberto Semenzin[6], Francesco Enrichi[1,2], Francesco Setti[1], Adolfo Speghini[7,2] and Marco Cristani[1,*]

[1]*Department of Engineering DIMI, University of Verona, Verona, Italy*
[2]*INSTM, RU of Verona, Verona, Italy*
[3]*Department of Earth Science, University of Torino, Torino, Italy*
[4]*Istituto Internazionale del Marmo, Italy.*
[5]*San Valentino Inerti, Verona, Italy*
[6]*Verona Stone District, Verona, Italy*
[7]*NRG-Department of Biotechnology, University of Verona, Verona, Italy*

Abstract

The field of material science is greatly benefitting from the introduction of artificial intelligence (AI) in research. In this context, we present the AI4Materials (AI4M) laboratory within the Intelligo network at the University of Verona, a multidisciplinary team for the development of AI in material science and related fields, with the final aim of speeding up research and make it more sustainable. The idea is to adopt a new research approach where AI and machine learning (ML) directly aid in determining crucial chemical properties and pinpointing bottlenecks in materials discovery and optimization, implementing them in real world cases thanks to a strong connection with industrial realities. While just in its infancy, the group is already involved in several projects, which are here presented and discussed in terms of challenges and results we reached so far, with the idea of sharing our vision on the current and future tasks.

Keywords

material science, machine learning, sustainability

1. Introduction

Materials science stands on the cusp of a revolution. As the demand intensifies for novel, high-performance, and environmentally sustainable materials, the limitations of conventional discovery methods are becoming increasingly apparent. Traditional research pipelines—dominated by time-consuming experimentation, limited generalizability, and slow feedback loops—are ill-suited to address the urgency and complexity of today's challenges. To meet the scale of global industrial and ecological demands, we must fundamentally rethink how we conduct scientific inquiry in this domain[1].

In the last years, thanks to the availability of large datasets and the increase in computational power, artificial intelligence (AI) has rapidly gained attention across several research fields, from which the material science community is not excluded. Innovative abilities, such as the discovery of novel compounds through thousands of structures screening or the optimization of existing ones on the basis of specific chemico-physical parameters, made AI-based techniques an appealing tool to foster research in this field [2]. For instance, when it comes to theoretical data, Barroso et al. [3] presented an Open Materials (OMat2024) database and pre-trained models, consisting of over 110 million quantum chemical calculations focused on structural and compositional diversity, while Zeni et

Ital-IA 2025: 5th National Conference on Artificial Intelligence, organized by CINI, June 23-24, 2025, Trieste, Italy
*Corresponding author.
eros.radicchi@univr.it (E. Radicchi); usman.syed@univr.it (U. Syed); federico.cunico@univr.it (F. Cunico); michele.cassetta@univr.it (M. Cassetta); uzair.khan@univr.it (U. Khan); paolo.marone@stonetechtraining.it (P. Marone); alberto.scarsini@inertisanvalentino.it (A. Scarsini); semenzin.filiberto@gmail.com (F. Semenzin); francesco.enrichi@univr.it (F. Enrichi); francesco.setti@univr.it (F. Setti); adolfo.speghini@univr.it (A. Speghini); marco.cristani@univr.it (M. Cristani)

al. [4] developed a generative model called MatterGen to design stable and diverse inorganic materials across the periodic table, which can be further fine-tuned to steer the generation towards a wide range of property constraints. But AI can also help in laboratory research, boosting efficiency in experimental design leveraging techniques such as supervised learning, active learning, reinforcement learning, and Bayesian optimization [5]. It should be then clear how machine-learning models are able to predict a wide range of diverse properties after training on both experimental measurements and high-throughput computational data. However, the majority of AI-based techniques need to work with vast and high-quality datasets, which is not always the case in fields such as materials science. Instead, we would like to focus our attention on all those situations where data are not abundant or their quality is not enough for a model to be properly trained: performing experiments in extreme conditions or with costly setup, working in small research laboratories where automation is poorly integrated or totally absent, dealing with data possibly covered by patents and industrial secrets, and for many emerging material classes still underexplored. In all these cases, data is frequently sparse and noisy, but support by AI techniques would be equally desirable both to speed up the research and to cut waste and costs associated with the intrinsic nature of practical experiments.

At AI4Materials (or, abbreviated, AI4M) — a research lab embedded within the Intelligo network at the University of Verona — we explicitly tackle this problem, proposing a scientific methodology that is inherently AI-native, simulation-driven, and physically grounded to gain the most from very small datasets. The team is multidisciplinary, with skills and knowledge ranging from physics and geology to chemistry and computer science, and is strongly oriented towards applied science, featuring a direct bond with industrial realities, which strengthens the link with the territory as well as technological transfer. Our mission is to demonstrate how materials research can be conducted under a unified framework where models, data, and lab processes are co-designed from the start for intelligent automation and discovery efficiency, with the final aim of optimizing experiments design and processes even at an industrial scale.

This paper introduces our newborn research group and its interests, showing at a glance the currently active projects and the directions we want to undertake. We believe that supporting materials science with innovative AI-based techniques will help the field to grow under the sign of sustainability.

2. Current projects

This section showcases current ongoing projects undertaken at the AI4M lab, each of which is presented in a distinct section, providing an exploration of the challenges we want to address, their underlying causes, and the AI-driven solutions we have developed or are currently developing.

2.1. Estimation of the growth of a particle using Physics-Informed Neural Networks

Accurate estimation of kinetic parameters governing phase transformations is a fundamental challenge in materials science, as it is strongly related to the final properties of a certain material. A widely adopted model to describe these transformations is the Avrami equation [6], with which data are fitted to obtain parameters that encapsulate mechanistic information about nucleation and growth processes. However, given common experimental limitations, these data are often sparse and noisy, making traditional curve-fitting methods unreliable and limiting the applicability of conventional data-driven models that lack physical interpretability [7].

To address this, we propose a methodology grounded in Physics-Informed Neural Networks (PINNs), which integrates physical laws directly into the training of neural models [8]. Instead of relying solely on data, PINNs encode domain knowledge—such as the Avrami equation for crystallization kinetics—into their loss functions, allowing them to remain physically consistent even when trained on limited or partially corrupted datasets. This dual reliance on data and governing equations makes PINNs an ideal tool for inferring hard-to-measure parameters from minimal experimental information.

Our framework also introduces a novel adaptive loss scaling strategy to automatically balance the contributions of the data fidelity and physics-informed terms during training. This circumvents the

manual tuning of hyperparameters, which is often a barrier to practical deployment of PINNs, and leads to improved training stability and generalization.

We demonstrate the potential of our method with an example from the literature that follows an Avrami-type kinetics, $i.e.$ the crystallization of silver nanowires [9]. Using just a handful of noisy time-series observations, our model successfully estimates the kinetic parameters with high fidelity. In comparative evaluations against classical optimization methods — including Levenberg-Marquardt, Trust Region Reflective algorithms, and global stochastic optimizers — our PINN consistently achieves comparable or superior fitting quality (coefficient of determination $R^2 > 0.998$) while showing stronger robustness in data-scarce regimes.

These results highlight the broader implications of our approach: by fusing data-driven flexibility with physical consistency, PINNs can serve as a generalizable framework for parameter estimation across a range of kinetic systems in materials science. Beyond crystallization, this paradigm may be extended to other domains such as reaction kinetics, diffusion processes, and solidification modeling, where data is limited and theory is rich.

2.2. Prediction of glass physical properties

Glass and glass-ceramics are central to numerous industrial and scientific domains, from structural and optical applications to volcanology. Glass is defined as *"[...] a nonequilibrium, non-crystalline state of matter that appears solid on a short time scale but continuously relaxes towards the liquid state"* [10] while glass-ceramics are inorganic, non-metallic materials prepared by controlled crystallization of glasses [11]. Two of the most critical parameters governing the behavior and processability of these materials are melt viscosity η (the liquid's ability to flow) and the glass transition temperature T_g (temperature at which $\eta = 10^{12}$ Pa $\cdot$ s). These properties determine the temperature range over which specific viscous responses occur during thermal treatments, such as strain, annealing, sintering, deformation, softening, and the working point—thereby enabling key processing operations like shaping, molding, fiber drawing, or extrusion [12]. In volcanology, they govern flow dynamics, transport behavior, eruption styles, and emplacement processes, all of which are critical for developing accurate volcanic hazard assessment scenarios [13].

Traditionally, the determination of T_g and η relies on experimental techniques such as differential scanning calorimetry and high-temperature rheometry. While accurate, these methods are often time-consuming, expensive, and dependent on specialized equipment, limiting their utility in high-throughput contexts or when sample quantities are constrained. Moreover, capturing the full temperature-dependent viscosity curve requires extensive measurements, which are not always practical or reproducible across different labs and conditions [12].

To address these limitations, we develop a ML-based framework capable of predicting T_g and η behavior from more accessible physical and compositional descriptors. Our approach leverages both static features—such as chemical composition and mechanical properties—and dynamic sequence data in the form of temperature-viscosity pairs. By employing a hybrid deep learning model, we effectively capture both temperature dependencies and complex correlations across different parental-glass properties.

This methodology provides a pathway toward more efficient, scalable, and accessible characterization of glass-forming systems. It enables the estimation and prediction of key thermophysical properties without direct measurement, paving the way for rapid material screening and informed design of amorphous systems across the silicate domain. More broadly, this work exemplifies the potential of AI-assisted modeling to bridge gaps between theory, experiment, and application in data-constrained regimes.

2.3. Optimization of marble waste mix design

The disposal of marble and stone-cutting sludge represents a significant environmental and economic burden, particularly in regions with a strong natural stone industry such as Northern Italy. Each

year, millions of tons of calcium-rich sludge—generated during the sawing, polishing, and shaping of marble—are discarded in landfills, incurring substantial costs and contributing to ecological degradation. In some areas, the cost of sludge disposal exceeds €5 million annually, highlighting the urgent need for alternative waste management strategies [14].

One promising route for valorizing this industrial byproduct is its incorporation into cementitious materials as a partial replacement for fine aggregates or binders. This process, known as mix design, has the dual benefit of minimizing waste and improving specific mechanical properties of the resulting composite. When properly formulated, marble sludge can enhance the compressive strength, workability, and durability of concrete mixtures. However, the current state of practice remains largely empirical. Mix proportions are typically derived through time-intensive trial-and-error methods, requiring repeated laboratory experiments that consume significant manpower and material resources [15, 16]. Such empirical approaches suffer from low scalability and poor generalizability: they are often tailored to specific local conditions or proprietary formulations, making them incompatible with evolving regulatory frameworks and varying material characteristics (for instance, marble sludge composition can vary significantly depending on the source stone), necessitating custom formulations for different batches. Together with industrial partners, such as Verona Stone District and Instituto Internazionale del Marmo (ISIM), we propose a simulation-driven framework that integrates ML and meta-learning to optimize mix designs for marble sludge reuse, see Figure 1. Our approach has been benchmarked on a public dataset with concrete properties from [16] using ML-guided simulations [17] to assess the feasibility of our adopted methodology and its further extension to the mix-design optimization challenge. Specifically, this dataset captures the relationships between input variables, such as water-to-binder ratio, slag and fly ash content, superplasticizer dosage, and curing conditions, and target properties like compressive strength, pull-off adhesion, and porosity.

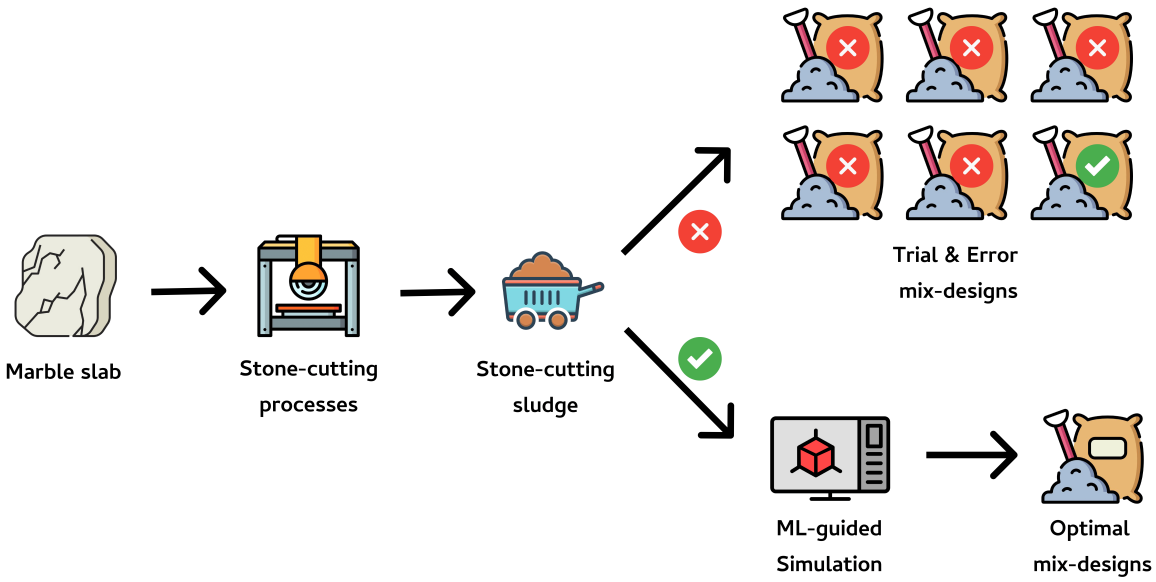

Figure 1: A visual overview of the proposed AI-driven framework for sustainable mix design using marble sludge.

Among these, porosity serves as a rapid and reliable proxy for material quality, strongly correlating with both mechanical performance and environmental resistance. In our preliminary experiments, we applied ensemble learning models such as XGBoost and Gradient Boosting with hyperparameter tuning via Bayesian Optimization, validating the simulation's capacity to guide experimental design as a promising strategy to also tackle the mix-design problem.

This work positions marble sludge not as a waste liability but as a high-value precursor in sustainable construction. By combining AI with domain expertise, we aim to establish a reproducible, scalable framework for eco-efficient material design, contributing to circular economy and supporting industrial

innovation in the stone sector.

3. Future directions

Building upon the methodologies and findings presented in this work, several promising avenues emerge for further development at the intersection of AI and materials science. Here, we present some of the future directions we want to undertake to carry on our projects:

1. From PINNs to PINOs (Physics-Informed Neural Operators):
While our current framework leverages PINNs to extract kinetic parameters from the Avrami equation, this formulation only assumes isothermal mechanisms. As a natural progression, we aim to develop a more generalizable approach for particle growth kinetics by employing *Physics-Informed Neural Operators* (PINOs). Unlike PINNs, PINOs are capable of learning solution operators that map function spaces to function spaces, enabling the modeling of nonlocal and highly nonlinear dynamics. This shift will allow us to learn governing growth laws directly from sparse experimental observations, offering broader applicability across iso and non-isothermal transformations.

2. Meta-Learning and PINNs for Robust Mix Design under Soft Physical Constraints:
In our work on sustainable concrete formulations using marble sludge, we demonstrated that meta-learning provides a data-efficient way to optimize mix designs across varying input conditions. A future direction is to hybridize this approach with PINNs, embedding domain-specific constraints (e.g., stress strain relationships) directly into the learning process. In particular, we propose combining *task augmentation* with *soft physical constraints* within a meta-learning framework. This would enable the model to generalize across different standards or precursor compositions while maintaining consistency with physical laws, improving both adaptability in industrial contexts.

3. Foundation Models for Materials Science with PINNS:
Recent advances in foundation models have shown strong potential in scientific domains by learning general representations from large, diverse datasets. These models has demonstrated remarkable performance on downstream tasks. Hence, we propose a foundational modeling approach tailored to materials science, where models pretrained on multi-source materials data are fine-tuned for specific tasks using PINNs. By embedding physical laws such as energy conservation and equilibrium into the training process, we aim to enable few-shot learning in complex materials systems with limited labeled data but rich mechanistic knowledge. This hybrid strategy promises to accelerate materials discovery and enhance model reliability.

4. Conclusions

This paper presents an overview of our contributions and the vision we are pursuing for advancing materials research through artificial intelligence. Our current methodology relies on the use of different state-of-the-art AI-based techniques, including PINNs as regularizers, Bayesian methods, and deep neural architectures depending on the particular problem we want to tackle. We briefly presented some of the projects we are currently working on, both to give insights on how we are applying these techniques and to show the kind of problems we want to address, which are strongly related to sustainability.

Looking forward, we intend to extend our methodology and develop an advanced framework that employs, for instance, neural operators to generalize kinetic models, meta-learning for data scarcity and foundational models for transfer knowledge and broader impact across diverse materials domains — all while remaining constrained by the physical realities that govern matter.

We hope that the ideas and frameworks presented here serve as both a practical contribution and a conceptual foundation for the next generation of AI-driven materials discovery.

Acknowledgements

This work was supported by the "FINAL LAP - FINALizzazione intelligente dei materiali LAPidei derivanti da taglio" project within the " Bando Ricerca e Sviluppo 2024".

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT-4o in order to: Drafting content. After using these tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] M. U. Arshad, Exploring the latest advances in materials science: Development of new materials with unique properties, Social Science Review Archives 2 (2024) 41–56.

[2] L. D., R. Nambiar, P. Kumar, 9 - the role of artificial intelligence in materials science: a review, in: A. Kumar, P. Kumar, V. Gambhir, R. C. Kuhad (Eds.), Advances in Sustainable Materials, Elsevier, 2025, pp. 209–228.

[3] L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick, Z. W. Ulissi, Open materials 2024 (omat24) inorganic materials dataset and models, arXiv preprint arXiv:2410.12771 (2024).

[4] C. Zeni, R. Pinsler, D. Zügner, A. Fowler, M. Horton, X. Fu, S. Shysheya, J. Crabbé, L. Sun, J. Smith, et al., Mattergen: a generative model for inorganic materials design, Nature 639 (2025) 624–632.

[5] L. Nolte, S. Tomforde, A helping hand: A survey about ai-driven experimental design for accelerating scientific research, Applied Sciences 15 (2025).

[6] K. Shirzad, C. Viney, A critical review on applications of the avrami equation beyond materials science, Journal of the Royal Society Interface 20 (2023) 20230242.

[7] L. Brigato, L. Iocchi, A close look at deep learning with small data, in: 2020 25th international conference on pattern recognition (ICPR), IEEE, 2021, pp. 2490–2497.

[8] M. Raissi, Deep hidden physics models: Deep learning of nonlinear partial differential equations, Journal of Machine Learning Research 19 (2018) 1–24.

[9] J. R. Sonawane, S. Patil, A. A. Kulkarni, Understanding the nucleation and growth kinetics of the microwave-assisted synthesis of silver nanowires, Industrial & Engineering Chemistry Research 62 (2023) 14199–14211.

[10] E. D. Zanotto, J. C. Mauro, The glassy state of matter: Its definition and ultimate fate, J. Non Cryst. Solids 471 (2017) 490–495.

[11] J. Deubener, M. Allix, M. J. Davis, A. Duran, T. Höche, T. Honma, T. Komatsu, S. Krüger, I. Mitra, R. Müller, S. Nakane, M. J. Pascual, E. D. Schmelzer, J. W.P. Zanotto, S. Zhou, Updated definition of glass-ceramics, Journal of Non-Crystalline Solids 5o1 (2018) 3–10.

[12] J. C. Mauro, Y. Yue, A. J. Ellison, P. K. Gupta, D. C. Allan, Viscosity of glass-forming liquids, Proceedings of the National Academy of Sciences 106 (2009) 19780–19784.

[13] C. G. Newhall, R. S. Punongbayan, The narrow margin of successful volcanic-risk mitigation, Monitoring and Mitigation of Volcano Hazards (1996) 490–495.

[14] L. Zichella, R. Bellopede, S. Spriano, P. Marini, Preliminary investigations on stone cutting sludge processing for a future recovery, Journal of cleaner production 178 (2018) 866–876.

[15] K. K. Alzboon, K. N. Mahasneh, Effect of using stone cutting waste on the compression strength and slump characteristics of concrete, International Journal of Environmental Science and Engineering 1 (2009) 167–172.

[16] J. Liu, F. Liu, C. Zheng, D. Zhou, L. Wang, Optimizing asphalt mix design through predicting effective asphalt content and absorbed asphalt content using machine learning, Construction and Building Materials 325 (2022) 126607.

[17] U. Syed, F. Cunico, U. Khan, E. Radicchi, F. Setti, A. Speghini, P. Marone, F. Semenzin, M. Cristani, Material synthesis through simulations guided by machine learning: a position paper, in: International Conference on Pattern Recognition, Springer, 2025, pp. 289–303.