AI-driven computational workflows for scalable innovative materials design

Stephan Roche[1,2,3], Dorye L.E. Córdoba[1], Andrei Tomut[2], Mariana Ionita[4], José-Hugo Garcia[1,2]

[1][Catalan Institute of Nanoscience and Nanotechnology (ICN2), CSIC & BIST, Campus UAB, 08193 Bellaterra (Barcelona), Spain] [2][Apeiron Intelligence, Barcelona, Spain] [3][ICREA, 08010 Barcelona, Spain]
[4][Faculty of Medical Engineering and Advanced Polymer Materials Group, National University of Science and Technology Politehnica Bucharest, Bucharest, Romania]

Correspondence to: Stephan Roche stephan.roche@icn2.cat

1. Introduction

Modern functional materials display coupled phenomena across many length and time scales, from quantum-level electronic structure and lattice dynamics to mesoscale disorder, interfaces, degradation pathways, and device-level responses. Delivering predictive insight therefore requires integrated multiscale workflows rather than isolated calculations. Here we present a pathway for highly efficient, AI-driven, reproducible workflows that connect accurate first-principles data to large-scale models and output observables that can be quantitatively compared with experiments. Beyond semiconductors, AI hardware, and energy-efficient technologies, the same framework is increasingly relevant to advanced (intelligent functional) biomaterials, where chemistry, microstructure, mechanics, transport, and biological performance must be co-optimized for applications such as regenerative scaffolds, neural interfaces, and adaptive soft matter.

2. Current and future challenges

Despite major progress, a universally adopted solution for materials workflows remains elusive. Key blockers are (i) fragmentation across codes and data formats, which hinders interoperability and reusability; (ii) limited standardization of provenance and metadata; and (iii) the high cost of generating high-quality training data for AI surrogates. These barriers become even more acute for complex matter such as biomaterials, where datasets are typically sparse, multimodal, and context dependent. In such systems, structure-function relationships must capture not only composition and disorder, but also porosity, hydration, degradation, interfacial charge transfer, manufacturability, and biocompatibility. Additional challenges include uncertainty quantification under limited data, interpretable models for safety-critical decisions, and the need to integrate lifecycle, regulatory, and sustainability constraints from the earliest design stages.

3. Approach and main contribution

We propose an AI-orchestrated workflow stack that couples: (1) first-principles electronic structure methods to generate accurate ground-truth data; (2) machine-learned surrogates, including deep-learning Hamiltonians, to emulate ab initio fidelity at reduced time and computational costs; (3) large-scale, linear-scaling quantum simulations such as LSQUANT (www.lsquant.org) to compute experimentally measurable observables and metrics on realistic system sizes; and (4) data-efficient active learning to minimize expensive calculations by selecting the most informative structures and conditions for retraining and validation. Recent advances in large language models enable a further layer: an AI agent that translates high-level natural-language objectives into verified multi-step simulations while preserving workflow logic and provenance to avoid black-box behavior. This is the typical enabling platform developed by our spinoff Apeiron Intelligence (www.apeironai.eu).

3.1 Related work and novelty

To extend this paradigm toward complex matter such as biomaterials, we shall incorporate multimodal data fusion and digital-twin thinking across microscopy, spectroscopy, mechanics, electrochemistry, and biological assays. This would enable the same workflow logic to connect atomistic or mesoscopic descriptors with application-relevant outputs such as degradation kinetics, pore-network metrics, interfacial transport, mechanical compliance, and surrogate predictors of biological response. In parallel, inverse design and closed-loop experimentation can prioritize candidate chemistries, architectures, and processing routes that jointly optimize function, robustness, and deployability. Our approach therefore combines workflow-native AI agents, active learning, LSQUANT-computed metrics, and a broader design space

that includes disordered inorganic materials together with functional intelligent biomaterials. This could create a scalable path from physics-grounded simulation to experiment-linked prediction for semiconductors, amorphous phases, interfaces, tissue-engineering scaffolds, adaptive hydrogels, and neural biointerfaces.

4. Outlook

The next step is to close the loop between simulation, automated characterization, building of digital twins of as-grown matter and extraction of design rules of materials under real-world constraints. For biomaterials, this means unifying computational descriptors with experimentally accessible readouts and with clinically or technologically meaningful metrics such as degradation, integration, reproducibility, and safe deployment. In this broader setting, AI becomes not only a computational accelerator but also a coordination layer for self-driving discovery, adaptive validation, and uncertainty-aware decision support across the full materials lifecycle.

Acknowledgments

ICN2 is supported by the Severo Ochoa Centres of Excellence programme (CEX2021-001214-S/MICIU/AEI/10.13039/501100011033). This work received funding from the European Union's Horizon Europe programme (ERC Executive Agency, grant 101078370 - AI4SPIN). Mariana Ionita acknowledges support from a grant from the Ministry of Research, Innovation, and Digitization, Romania's National Recovery and Resilience Plan, under the project titled 'Advanced & Personalized Solutions for Bone Regeneration and Complications Associated with Multiple Myeloma', with reference CF 213/29.11.2022 and contract number 760093/23.05.2023.