PREMISE: Enabling seamless data exchange for future autonomous labs

E. Bainglass[*1,4], F. Lopes[2], N. Plainpan[2], A. Laskowski[3], M.C. Danaila[3], A. Meier[3], J. Fuentes[3], A. Yakutovich[2], C. Barillari[3], B. Rinn[3], C. Pignedoli[2], C. Battaglia[2], and G. Pizzi[1,4]

[1] Paul Scherrer Institute (PSI), Forschungsstrasse 111, 5232 Villigen PSI, Switzerland
[2] Eidgenössische Materialprüfungs- und Forschungsanstalt (Empa), Ueberlandstrasse 129, 8600 Dübendorf, Switzerland
[3] ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
[4] National Centre for Computational Design and Discovery of Novel Materials (MARVEL)

[*]edan.bainglass@psi.ch

Key Words: *Interoperability, Autonomous labs*

Abstract

The discovery of new materials is essential to addressing pressing societal challenges. To accelerate discovery, modern technologies, such as artificial intelligence, robotics, and supercomputing, can be integrated into autonomous research laboratories operating in closed-loop feedback involving experiment and simulation. However, a key challenge lies in the seamless exchange, tracking, and provenance of open research data (ORD) across experimental and simulation platforms, which often employ different formats, standards, and APIs. The PREMISE project[1] of the ETH Board's ORD program[2] is addressing these challenges primarily within the materials sciences by establishing best practices, standards, and protocols for ORD exchange adhering to FAIR principles. Though tested internally between the AiiDA[3] workflow management system (WFMS) and the openBIS[4] ELN/LIMS platform, the project aims more broadly to deploy platform-agnostic APIs leveraging semantic annotation via the RO-Crate[5] and JSON-LD[6] specifications. PREMISE also sponsors the MADICES[7] workshop series bringing together ontology experts and RDM platform representatives to drive the adoption of standards and protocols. In this presentation, we will highlight recent developments towards enabling seamless integration of ORD from automated and/or simulation-assisted experiments.

References

[1] https://ord-premise.org/
[2] https://ethrat.ch/en/eth-domain/open-research-data/
[3] Sebastiaan. P. Huber, et al., AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance, Scientific Data 7, 300 (2020)
[4] C. Barillari, et. al., openBIS ELN-LIMS: an open-source database for academic laboratories, Bioinformatics 32, 4, (2016)
[5] https://www.researchobject.org/ro-crate/
[6] https://json-ld.org/
[7] https://madices.github.io/