Publishing reproducible results supported by FAIR data

Abraham Nieva de la Hidalga[1,2]*, Leandro Liborio[3], Patrick Austin[4], Alejandra González-Beltrán[4]

[1]UK Catalysis Hub, Research Complex at Harwell, Rutherford Appleton Laboratory, R92, Harwell Oxford Oxfordshire OX11 0FA,

[2] Cardiff University, School of Chemistry Cardiff University, Translational Research Hub Building, Room 4.08, Maindy Road, Cardiff, CF24 4HQ

[3]Theoretical and Computational Physics Group, SCD, STFC, Rutherford Appleton Laboratory, R89, Harwell Campus, Didcot, OX11 0QX, UK

[4]Data and Software Engineering Group, SCD, STFC Rutherford Appleton Laboratory, R89, Harwell Campus, Didcot, OX11 0QX, UK

corresponding: nievadelahidalgaa@cardiff.ac.uk

CatalysisHub

Research Complex at Harwell,
Rutherford Appleton Laboratory,
STFC Harwell Campus, Oxfordshire

Reproducibility requires data to be Findable, Accessible, Reusable and Interoperable (FAIR)

FAIR data includes:

- Explicit data-result mappings (Data 'X' is used to produce Figure 'A')
- Clear identification of resources required (crystal structures, models)
- Clearly stated processing parameters (E0, energy ranges)

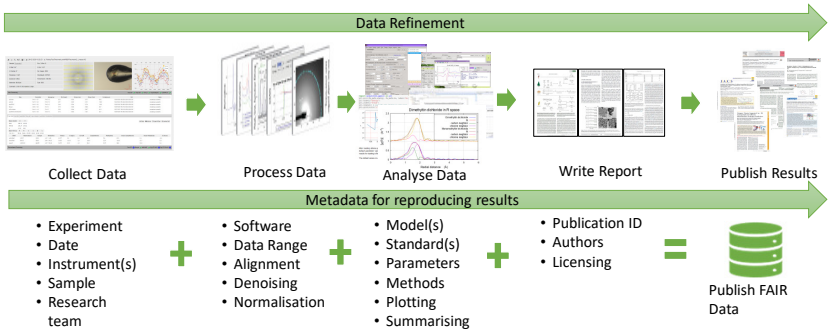
PROBLEM: Complex computational studies require increasing effort and time to progress in the scale from repeatability to runnability, reproducibility, and replicability.

10.1016/j.envsoft.2020.104753

Proposal: develop or adopt tool sets which can be used to process and analyze data while producing the required mapping data with no extra effort.

XAS Processing and Analysis example

python™



X-Ray Larch based XAS processing examples

python™ Larch Data Analysis Tools for X-ray Spectroscopy

jupyter FUSION MLProvLab

Python, X-Ray Larch, Jupyter, MLProvLab

- Desktop based
- Interactive
- Small-Medium scale spectra analysis (from ex-situ or in-situ experiments)

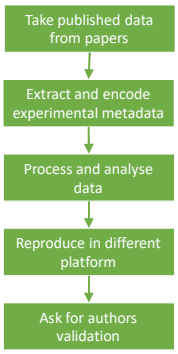
python™ Larch Data Analysis Tools for X-ray Spectroscopy

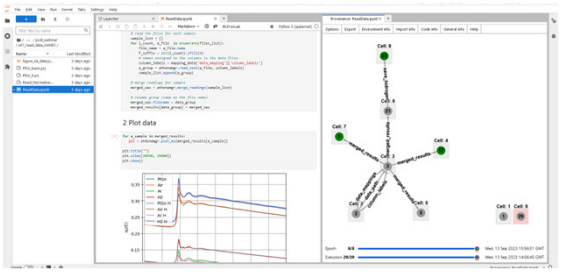
Galaxy PROJECT

Python, X-Ray Larch, Galaxy

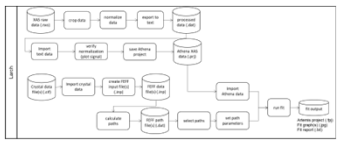
- Web Based
- No setup required
- Mediul-Large scale spectra analyses (from in-situ or operando experimets)

Testing the theory

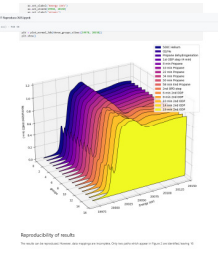




End to end EXAFS processing and analysis with Larch



Catalysis Data Infrastructure (CDI) is a catalogue of UKCH publications and their supporting research data objects.



Catalysis Research Workbench (CRW) will be a catalogue of tools and resources for analysis and processing of research data objects

PSDI PHYSICAL SCIENCES DATA INFRASTRUCTURE Pathfinder

Testing how the CDI and the CRW can be integrated and make use of the services and resources to be provided by the PSDI.

Two interdependent topics were chosen:

- **FAIR Digital Objects** packaged reproducible results
- Scientific Workflows for processing and analysis

CatalysisHub

UKRI Research Complex at Harwell

CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD

UKRI Science and Technology Facilities Council

ada lovelace centre