



Authors

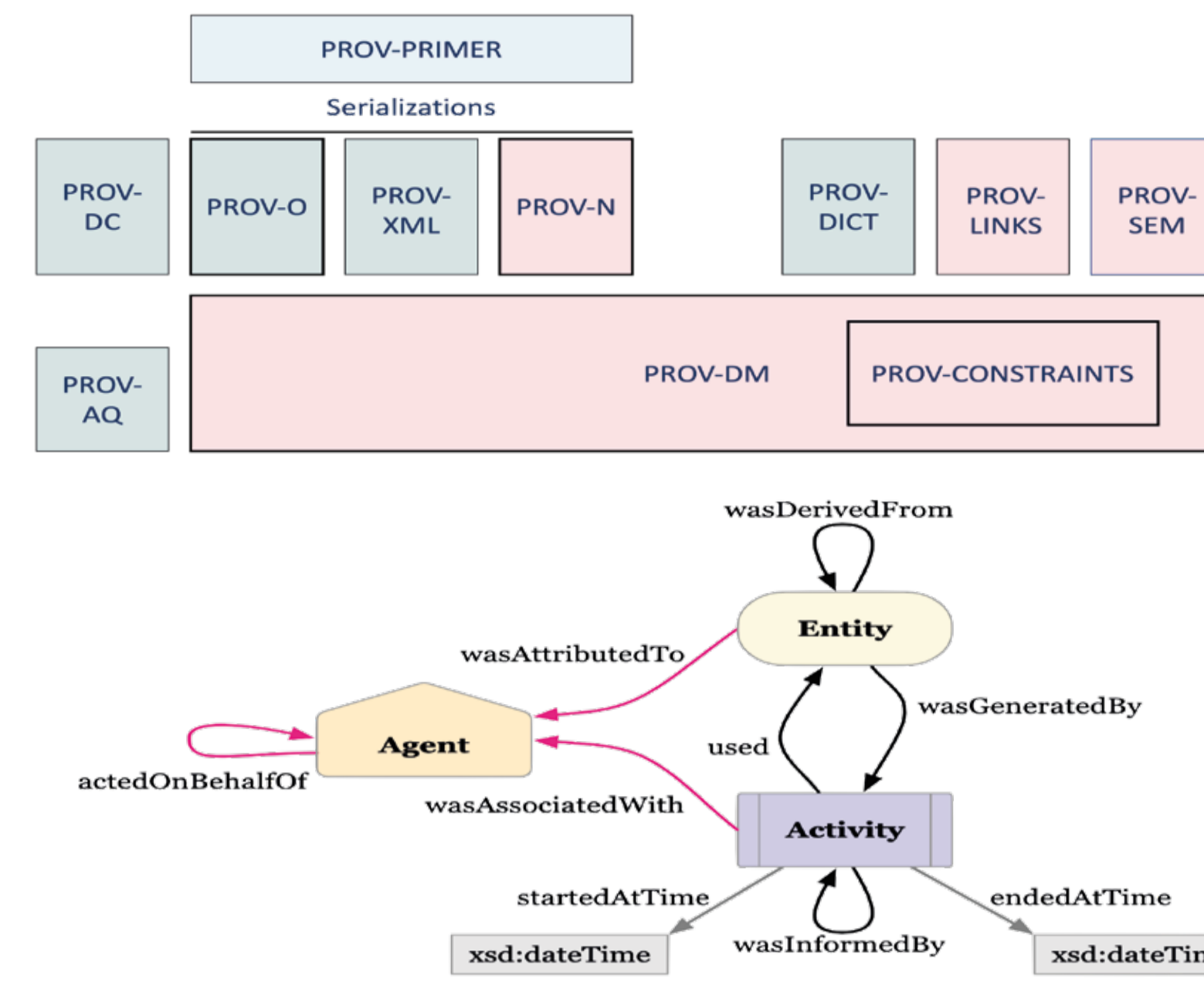
Authors: Elda Osmenaj¹, Fabrizio Antonio², Marco Kulüke³, Tommaso Rodani⁴, Sandro Fiore⁵, Donatello Elia², Alessandro D'Anca², Mirco Panighe¹, Alberto Cazzaniga⁴, Pablo de Andres⁶, Africh Cristina¹, Stefano Cozzini⁴
¹National Research Council-IOM Materials Foundry (CNR-IOM), Italy; ²Advanced Scientific Computing Division, Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy; ³Deutsches Klimarechenzentrum GmbH (DKRZ), Germany; ⁴Area Science Park (AREA), Italy; ⁵Department of Information Engineering and Computer Science, University of Trento, Italy; ⁶Fraunhofer Institute for Mechanics of Materials IWM

Introduction

In the context of the EOSC-Pillar project, Work Package 6 (*EOSC in action: Use cases and community-driven pilots*) delivers use cases to analyze different tools and services used for the "FAIRification". Among them, Use Case 1 (*Defining procedures and services to enforce data provenance for thematic communities and beyond*) aims at elaborating cross-domain, FAIR-oriented procedures and recommendations to enforce data provenance in two scientific domains: **Materials Science/Nanoscience and Climate Science**.

Description

Description
 Due to the increasing complexity of data analysis workflows, **provenance management** is a key component to support the re-use of scientific data. The **W3C PROV** standard is considered for tracking processes and responsibilities as well as describing provenance structures within complex experiments and large workflows.



The **PROV Family of Documents** defines a data model (**PROV-DM**) along with the corresponding serializations (**PROV-N**, as a notation aimed at human consumption, **PROV-O** as the ontology for mapping the PROV data model to RDF, **PROV-XML** as XML schema for the PROV data model) to enable the inter-operable interchange of provenance information in heterogeneous environments.

Materials science/ Nanoscience case study

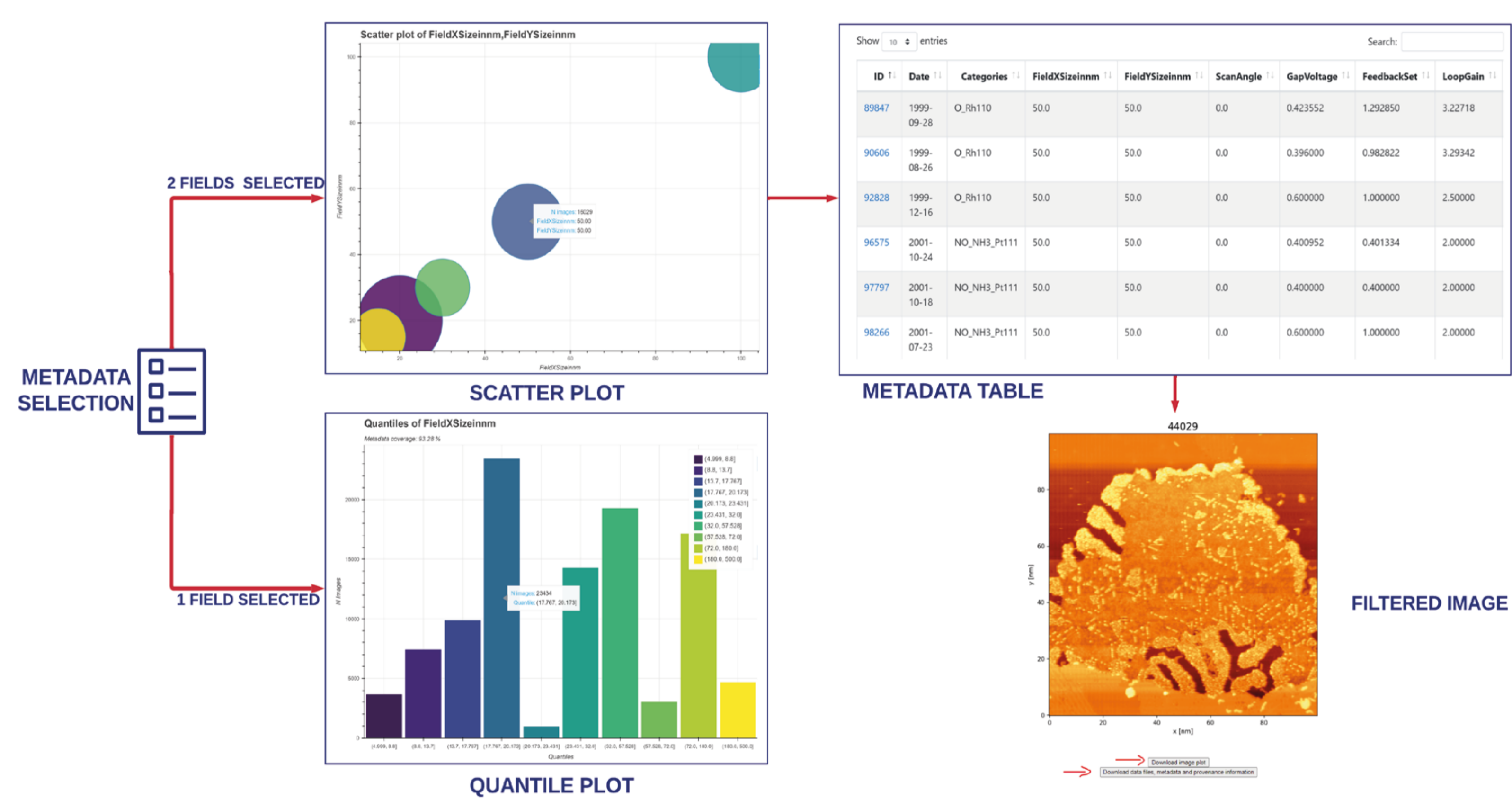
Main objective

Provide **guidelines** and **recommendations** on how to manage the important aspect of **data provenance** for the wide community of the **NFFA-Europe PILOT project (NEP)**.

Description

Development of data management practices and services for making **FAIR** compliant a scientific archive of **Scanning Tunneling Microscopy (STM)** images.

The web service workflow on the TriDAS website



Results

- ★ Application of the **W3C PROV** standard to describe provenance metadata
- ★ Provision of services and tools designed to improve the value of the **STM dataset**
- ★ Implementation of different aspects of **FAIR principles**

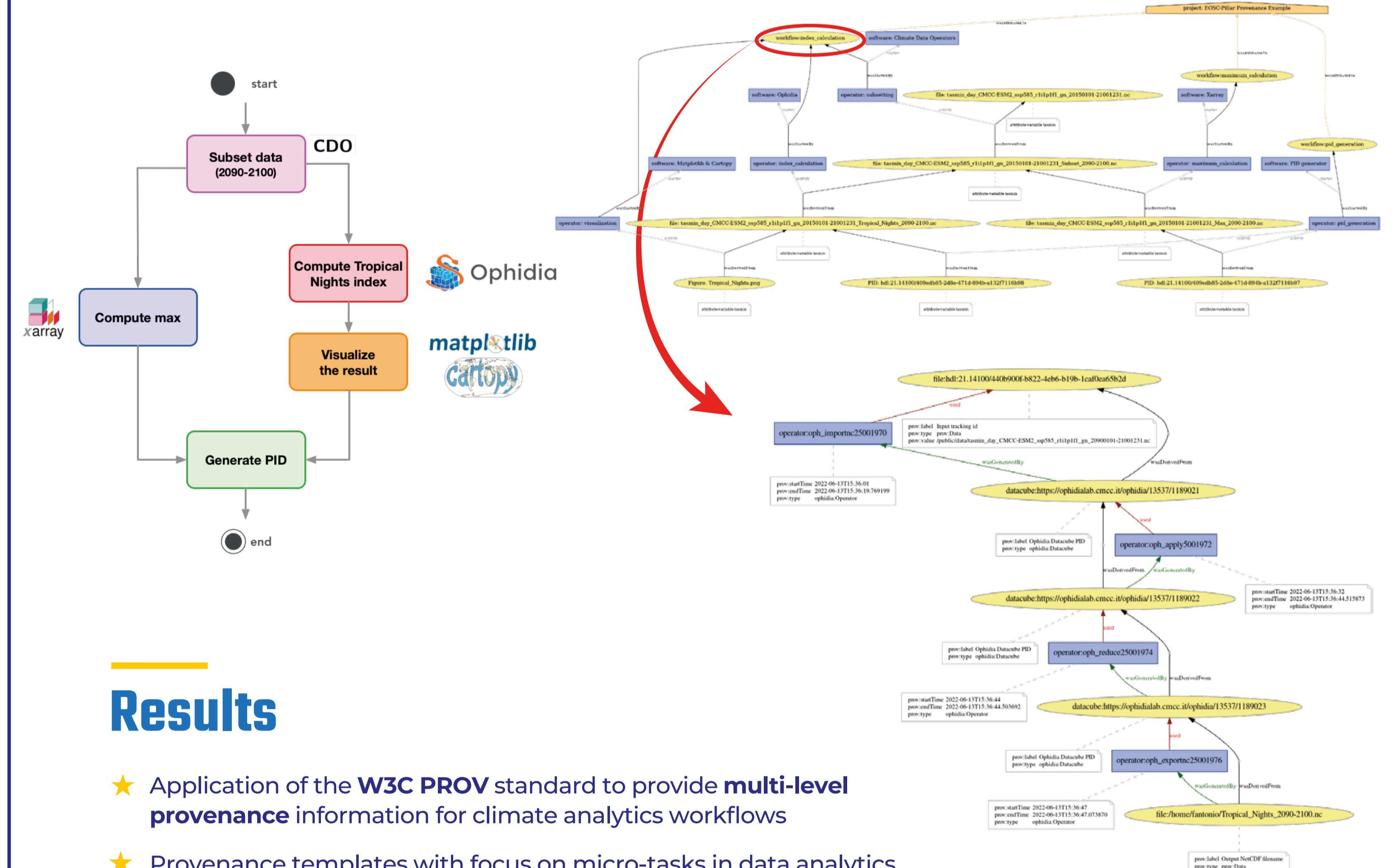
Climate Science case study

Motivation

Climate research makes use of lots of data coming from the modelling and observational climate communities. In this domain, provenance management plays a key role both for **numerical end-to-end simulations** at data center level and in the **inner data analytics workflows**.

Main objective

- Address **provenance tracking** at two different levels:
- ★ understand the whole end-to-end scientific workflow including a proper reference to input and output datasets via PIDs (**first-level**);
 - ★ drill-down into specific first-level tasks (workflows of hundreds or thousands analytics operators) to get more detailed information (**second-level**).



Results

- ★ Application of the **W3C PROV** standard to provide **multi-level provenance** information for climate analytics workflows
- ★ Provenance templates with focus on micro-tasks in data analytics

Conclusion

The proposed approach enables **shareability** and **re-usability** of workflows and results across the community. This allows an assessment of the quality, reliability and trustworthiness of the data, thus promoting **Open Science** and fostering new opportunities for **research & collaborations**.

More details



UCI page on the EOSC-Pillar website:
<https://www.eosc-pillar.eu/use-cases/defining-procedures-services-enforce-data-provenance-thematic-communities>



Dataset of Scanning Tunneling Microscopy (STM) images of graphene on nickel. <https://doi.org/10.5281/zenodo.5799773>



UC6.1 Factsheet:
<https://zenodo.org/record/6725047#.YzsiTdjP1nI>



Tommaso Rodani, Elda Osmenaj, Alberto Cazzaniga, Mirco Panighe, Cristina Africh, Stefano Cozzini; Towards the FAIRification of Scanning Tunneling Microscopy Images. Data Intelligence 2022;
 doi: https://doi.org/10.1162/dint_a_00164



Guidelines for researchers:
<https://zenodo.org/record/6827750#.YzVqoC8QPyg>



Partners



EOSC-Pillar has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 857650.

Contact

Elda Osmenaj
 National Research Council-IOM Materials Foundry (CNR-IOM), Italy
 E-mail: osmenaj@iom.cnr.it