Contribution ID: **92** Type: **Presentation**

Documenting Heritage Science: A CIDOC CRM-based System for Modelling Scientific Data

Tuesday, 9 October 2018 17:00 (15 minutes)

The paper presents a complete system for documenting scientific data produced in heritage sciences, based on a data model intended to generate valuable information and suitable metadata for them to be stored, accessed, queried, shared and reused in various contexts and research scenarios.

The system is built around the concept of a general meta-model, flexible enough to provide descriptions, in a formal language, of all the entities and issues encountered documenting heritage sciences results. It is inspired by CIDOC CRM principles for data modelling and maintains a full compatibility with CIDOC CRM and its extensions, especially CRMsci, CRMdig and CRMpe. The use of a wide set of thesauri and vocabularies for the standard and unambiguous description of all the entities will guarantee internal coherence at data level. Thus, our system is capable of identifying and modelling physical and digital objects, events, activities and actors, i.e. people and teams involved in the various research events and of providing straightforward connection and interoperability with the general documentation of cultural heritage, tightly linking scientific analyses to their heritage context.

The metadata model supported by our system will also stand at the very foundations of DIGILAB, a digital infrastructure developed by the E-RIHS European initiative to facilitate virtual access to tools, services and data for heritage research. DIGILAB infrastructure will rely on a network of federated repositories and will enable finding and accessing data through an advanced semantic search system operating on a registry containing metadata describing individual datasets. Thus, our data model is designed to make DIGILAB compliant with the EU policies and strategies concerning scientific data, including the FAIR data principles, the Open Research Data policy, and the EOSC strategy. It will guarantee data interoperability and will foster re-use of information and services to process the data according to specific research questions and use requirements.

First tests have been carried out on datasets resulting from various scientific analyses carried out by different research institutions, including the Italian National Council of Researches (CNR), the Istituto Superiore per la Conservazione ed il Restauro (ISCR), the Opificio delle Pietre Dure (OPD) and the National Institute of Nuclear Physics (INFN). A specific subset of information derived from the activity of the INFN-CHNet network (Cultural Heritage Network of the Italian National Institute for Nuclear Physics) are reported in this paper. The heterogeneity of the network analytical techniques examined and encoded, has ensured a good test bench for the developed model, proving its effectiveness and delineating a solid path for its future developments.

The encoding of scientific information by means of our system demonstrates the validity of this approach to different cases and offer an overview of the whole model and of how information encoded by means of its classes and properties will benefit the implementation of the DIGILAB infrastructure.

Type of abstract

Presentation

Summary

The paper presents a complete system for documenting scientific data produced in heritage sciences, based on a data model intended to generate valuable information and suitable metadata for them to be stored, accessed, queried, shared and reused in various contexts and research scenarios. The system is inspired by CIDOC CRM principles and is fully compatibility with the CIDOC CRM ontology and its extensions.The model is intended to benefit the implementation of the DIGILAB, a digital infrastructure developed by E-RIHS European initiative to facilitate virtual access to tools, services and data for heritage research to scientific community.

Primary author: FELICETTI, Achille (PIN S.c.R.L.)

Co-author: Dr CASTELLI, Lisa (INFN)

Presenter: Dr CASTELLI, Lisa (INFN)

Session Classification: Data Management Services

Track Classification: Area 2. Data science and skills