Global Information Infrastructures for Science & Cultural Heritage: The Interoperability Challenge, Costantino Thanos, Institute of Information Science & Technologies, National Research Council of Italy (CNR-ISTI)

As technology becomes more far-reaching and connected, interoperability is crucial when building heterogeneous interconnected information infrastructures. Interoperability is a multi-layered and context-specific concept, encompassing different levels along a multidimensional spectrum. At one end of the spectrum we have data and metadata interoperability while at the other end we have organisational, legal and policy interoperability.

Building global heterogeneous interconnected information infrastructures requires addressing all dimensions of interoperability including content, user, functionality, architecture, quality, and policy.

In addition, for each of these dimensions, diverse interoperability levels must be considered: technical/basic, syntactic, semantic, functional, operational, behavioural, secure, organisational/business, temporal, and language interoperability. These levels form dependent layers: operational interoperability is only possible if semantic interoperability is ensured; semantic interoperability is only possible if syntactic interoperability is ensured; and so on.

There are two main impediments to achieving interoperability:

- The heterogeneity of the exchanged information which covers all types of syntactic, structural, and semantic diversities among systems used to modelling information.
- The inconsistency between the use of the information as intended by its originator and the intended exploitation of it by the recipient.

The main concept enabling interoperability is mediation. This concept has its roots in tackling the integration of heterogeneous data sources and has been used to address a number of dimensions of heterogeneity spanning terminology, representation format, transfer protocols, semantics, and so forth. The mediation concept is implemented by a mediator, which is a software device that supports a mediation schema capturing user requirements, and an intermediation function between this schema and the distributed information object sources. A key feature that characterises a mediation process is the kind of intermediation function implemented by a mediator. There are two main functions: Mapping and Matching.

Mapping refers to how information object structures, properties, relationships are mapped from one representation scheme to another one, equivalent from the semantic point of view. Matching refers to the action of verifying whether two strings/patterns match, or whether semantically heterogeneous information objects match. There are several approaches to implementing the intermediation function, of particular relevance the approaches based on standards and ontologies. In order to effectively implement the intermediation function exchanged information objects should be complemented with contextual, provenance/lineage, and quality information.

DL.org addresses interoperability problems and solutions from several perspectives: content, user, functionality, policy, quality, and architecture, offering a forum where stakeholders, programme owners, researchers and implementers can meet for the purpose of exchanging experiences and formulating guidelines for addressing interoperability. DL.org workshops and presentations deliver insights into relevant interoperability dimensions and levels for a number of important application domains, spanning Digital Libraries, Cultural Heritage, and eScience, presenting guidelines to determine the most effective way to achieve interoperability.

Contributing projects:

- DL.org addressing the problem of Digital Library Interoperability.
- GRDI2020 addressing enabling technologies for Global Data infrastructures.
- DC-NET addressing Cultural Heritage infrastructures.

Funded by the EC under the 7th Framework FProgramme7, the DL.org Co-ordination Action addresses two important and closely related Digital Library issues: (a) strengthening the modelling foundations of the field through consolidation and enhancement of the DELOS DL Reference Model and (b) identifying requirements, solutions, and future challenges for achieving Digital Library Interoperability. The Reference Model has defined the main concepts characterising Digital Libraries and Digital Library Systems, that is, architecture, content, functionality, quality, policy, and user. Given the widely distributed nature of future Digital Libraries, heterogeneity is expected to be the norm with respect to all these concepts. Therefore, techniques for interoperability are crucial in reconciling different approaches in such systems. DL.org is undertaking a comprehensive analysis of interoperability in the domains of all six concepts mentioned above through a critical review of the current situation and of emerging trends. The goal is to identify techniques, methods, and approaches for DL interoperability based on the requirements of existing systems. These objectives are being pursued with the support and active contributions of the international DL research community, specifically through six Thematic Working Groups (one for each DL concept) and a Liaison Group, ensuring broad participation of the community. On-going activities are bringing into sharp relief solutions, best practices and shared standards for interoperability and are identifying future research challenges in the field.

Interoperability, with a special focus for the level of programmes and planning of investment, is a key issue for the DC-NET project, where eight Ministries of Culture are engaged in preparing a plan of joint activities for the implementation of the European digital cultural heritage e-Infrastructure. To this end, DC-NET co-ordinates national and European programmes to enable the use and development of e-Infrastructures for pan-European cultural heritage research. It defines the interface between cultural heritage and e-Infrastructures nationally and internationally, with the ultimate aim of fostering tangible co-operation in the long term.

GRDI2020, a Global Vision for Global Research Data Infrastructures, is a Coordination Action funded under the 7th Framework Programme by the GÉANT and Infrastructure Unit of the European Commission. GRDI2020 is defining a framework of technological, organisational and policy recommendations to guide the development of a GRDI ecosystem by mobilising user communities, international initiatives, worldwide experts and policy makers. GRDI2020 coordinates two Working Groups addressing technological and organisational issues comprising experts with a proven track record in these fields. These focused activities are underpinned by insights delivered by an Advisory Board comprising internationally renowned experts. Additionally, GRDI2020 supports the EC's High Level Expert Group chartered with defining a 2030 vision for scientific data infrastructures. The final output will be a Roadmap addressing application, technological and organisational aspects involved in the development of a sustainable global scientific data infrastructure.

Major Stakeholder Communities

Major stakeholders contribute to the activities of the host projects, including European initiatives, such as Europeana & several of its related projects; DARIAH; DRIVER & D4Science; Planets; SHAMAN and EU Ministries of Culture. From the perspective of DC-NET, the strategic alliance is an opportunity to deliberate how objectives and ambitions of the cultural heritage sector can be co-ordinated in an interoperable way while moving towards ICT research networks as well as the manner in which AHSS actors can and should engage with national and European e-Infrastructures.