

Title: eSciDoc - a service infrastructure for management of Cultural Heritage content

Type of Paper: PROJECT PAPER

The eSciDoc project (<http://www.escidoc.org>), developed by Max-Planck Digital Library (MPDL) and FIZ Karlsruhe, provides an open source infrastructure for the scientists to store, preserve, retrieve and work with research-based data. Since multiple and fast-evolving requirements demand a broad set of reusable, generic and flexible components, it is designed and developed as a service-oriented architecture. In addition, community-specific solutions built around the services allow user- and research-focused working environments. The heterogeneity of research questions, tools, workflows and primary data, as well as traditional forms of publications required us to focus on supporting multiple content models and descriptive metadata formats, together with common functionalities such as persistent identification, adequate versioning and management of primary data, aggregation of data, annotations, access control etc. The content models are formal representations of discipline-specific data models (e.g. integrated image+text view of primary sources, precisely documented collections of images) and their respective validation rules. Relations between content objects (e.g. organisation within collection, kinship relations) are expressed by semantic relations, defined with respective relation ontologies, which are applicable to any kind of content managed in the eSciDoc repository. The underlying Fedora repository, enriched with the eSciDoc core services layer, allows for management and structuring of data on different levels of granularity - from basic items to complex aggregations of data. Consequently, the eSciDoc infrastructure and its services can be used for cooperative working environments and management of content, as well as a mere archive solution for content managed in external systems. In addition, some services can be integrated by external applications; independent from the overall eSciDoc infrastructure.

In the course of the DARIAH project, the MPDL and its partners aim at building a pan-european research infrastructure based on Fedora/eSciDoc. This infrastructure will enable transparent access to and management of digital content across all participating museums, libraries and other institutions in the Arts and Humanities sector.

We would like to present some of the currently implemented solutions and our experiences with diverse content (publications, digitised books, collection of images used for research of human development) and provide an insight of how such architecture has been oriented towards preservation of the cultural heritage.