Digital Repository Infrastructure Vision for European Research - DRIVER –[1]

Sophia Jones[1], Paolo Manghi[2],

[1] Greenfield Medical Library, University of Nottingham,
Nottingham, United Kingdom
sophia.jones@nottingham.ac.uk

[2] Istituto di Scienza e Tecnologie dell'Informazione,
Consiglio Nazionale delle Ricerche
Pisa, Italy
paolo.manghi@isti.cnr.it

Abstract. This paper describes the DRIVER European Project mission, in terms of its organizational and technological goals and results.

Keywords: repository infrastructure, digital libraries, European project, Open Access

1 Introduction

The EU-funded project DRIVER ("Digital Repository Infrastructure Vision for European Research")[2] has two main aims: (*i*) setting-up a European *Confederation* for advocating and promoting EC Open Access mandates across European researchers and institutions, and (*ii*) enabling a *technical infrastructure* of European Institutional Repositories aggregating and making accessible Open Access content throughout Europe.

[1] This paper was originally published as an article in BELIEF Magazine 2009
[2] www.driver-community.eu

1.1 The DRIVER Confederation

The DRIVER confederation is an organization of people working to achieve a formal establishment of an (European) digital repository community. The Confederation mirrors strategic alliances that move towards a global, interoperable, trusted, long-term repository infrastructure for which DRIVER has built the nucleus in Europe. It aims to encourage a combined effort of repository development between federations within a network of content providers. The Confederation partners represent European and international repository communities, subject based communities, repository system providers, service providers, as well as political, research, and funding organisations who share the DRIVER vision to allow all research institutions in Europe and worldwide to make all their research publications openly accessible through institutional repositories. DRIVER liaises with institutions and initiatives from the majority of European countries, the U.S., Canada, Latin America, China, Japan, India and Africa.

Closely related to the theme of interoperability and specifically in relation to this project, are the DRIVER Guidelines, as they create a common ground by achieving interoperability on two layers: (*i*) syntactical (use of OAI-PMH and OAI_DC), and (*ii*) semantic (use of vocabularies). The data in the technical infrastructure is based on locally hosted resources that are collected in digital repositories and harvested and aggregated by DRIVER. In order to ensure a high quality of aggregation, the DRIVER Guidelines have been developed to make it possible to harmonise and validate the data. DRIVER makes its aggregated data available for re-use via OAI-PMH to all partners in the DRIVER network of content providers whilst respecting the provenance of resources by "branding" them with information of the local repository. The DRIVER Guidelines provide orientation for managers of new repositories to define their local data-management policies, for managers of existing repositories to take steps towards improved services and for developers of repository platforms to add supportive functionalities in future versions. By following the Guidelines repositories can become part of the DRIVER network and can re-use DRIVER data for the development of local services. In essence then, the DRIVER Guidelines assist repository managers to make their material more widely available. Interoperability in the sense

of the DRIVER Guidelines means standardised metadata of the harvested records, based on the use of standards.

1.2 The DRIVER Technical Infrastructure

An important outcome of the DRIVER project is *D-Net*,[3] a software toolkit capable of enabling a running environment where data and service providers can (*i*) find the tools to aggregate heterogeneous OAI-PMH compliant Institutional Repositories into uniform shared Information Spaces, and (*ii*) dynamically build, modify and customize their Digital Library (DL) applications to operate over such Spaces. The technology supports a Service-oriented e-infrastructure, where distributed and shared resources are implemented as standard Web Services and applications consist of sets of interacting services. The current implementation of D-Net offers functionality services required to build distributed aggregation systems and DL end-user applications. Aggregation systems enable the construction of uniform Information Spaces of metadata records to be harvested from possibly heterogeneous Institutional Repositories. Important in this context are Store Services, Index Services and Aggregation Services, which offer advanced tools for OAI-PMH harvesting, cleaning and integrating metadata records according to target metadata record formats. The resulting Information Spaces can then be accessed via an arbitrary number of DL applications built by service providers by combining D-Net services such as Recommendation, Collection, Browsing, User Interfaces and others.

Key features of D-Net are the *scalability* and the *openness* of its infrastructure instances. Specifically, an instance of the DRIVER infrastructure can scale up to arbitrary numbers of service instances, applications and organizations while the underlying application framework is open to the introduction of services providing new functionality so as to extend the D-Net toolkit.

Since July 2008, the DRIVER project maintains a running instance of D-Net which hosts one main aggregation system, integrating Open

[3] http://www.driver-repository.eu/D-NET_release: developed at Istituto di Scienza e Tecnologie dell'Informazione, CNR (Pisa, IT), ICM (Warsaw, PL), Department of Informatics and Telecommunications of the University of Athens (Athens, GR) and University of Bielefeld Library (Bielefeld, DE).

Access metadata records from a growing number of European Institutional Repositories. At present, the infrastructure runs 36 services distributed over 9 partner sites; as a result of the Confederation efforts, the resulting Information Space numbers 1,000,000+ records out of 200+ repositories across 27 countries, and the number of repositories willing to join is still growing. Currently, the space is accessed by three DL applications: the Belgium national repository portal, offering search over the Belgium Repository Federation subset; Recolecta national repository portal, offering search on the Spanish Repository Federation subset; and the main DRIVER portal, providing access and advanced functionality over the whole space.