

DRIVER: Digital Repository Infrastructure Vision for European Research

Donatella Castelli

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), Pisa, CNR

The DRIVER project (Digital Repository Infrastructure Vision for European Research - <http://www.driver-repository.eu/>) is funded by the European Commission under the auspices of the "Research Infrastructure" unit, and responds to the vision that any form of scientific-content resource, including scientific/technical reports, research articles, experimental or observational data, rich media and other digital objects should be openly, representing accessible to both researchers and the general public. Its partnership involves 13 European institutions, including the major national repository initiatives and technological partners.

DRIVER aims to provide a unified approach to manage the Digital Repository landscape by building an infrastructure that allows European Research Institutions to share content and functionality.

With the growth of repositories, many valuable services have been established to search and retrieve bibliographic records (metadata). However, building individual repositories requires high installation and maintenance costs, and items held in repositories are sometimes hidden behind several intermediate pages, obscured by authorization procedures, not fully presented, or not retrievable at all. In aid of furthering, improving and developing the current situation, DRIVER's primary objective is to establish a cohesive, pan-European infrastructure of Digital Repositories, providing its users with sophisticated functionality services on potentially all European and world-wide Digital Repositories.

The project's first important result has been the development of an advanced-quality test-bed infrastructure, based on a distributed service-oriented architecture enabling enhanced interoperability of data and service-providers offering fundamental generic functionality. Its design is characterized by great flexibility and scalability, so that its progressive expansion can move from building individual repositories or repository clusters, and proceed repeating such an aggregations, thus building a repository structure hierarchically organized according to different criteria.

In its initial phase, the Project has also conducted studies on the state of the Digital Repository landscape to identify its multifaceted aspects, with respect to different countries, different types of resources, different technological platforms, different metadata policies etc. As repository networks critically depend on interoperability in terms of technology and content provision, the Project has also defined joint policies in order to provide guidance for technicians and managers and also to direct future developers to adopt commonly accepted solutions thus favouring interoperability.

Presently, DRIVER infrastructure manages 117 institutional repositories, but this number is expected to rapidly grow in the next future. The project's second phase will see the construction of a production-quality infrastructure delivering services 24 hours a day, 7 days a week. Technological developments will focus on the enhancement of services for specific communities and the management of datasets and multimedia and multi-type compound objects.

DRIVER: Digital Repository Infrastructure Vision for European Research

Donatella Castelli

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), Pisa, CNR

DRIVER è l'acronimo di un progetto che vuole realizzare un'infrastruttura per permettere l'accesso aperto agli archivi istituzionali delle organizzazioni della ricerca europea (Digital Repository Infrastructure Vision for European Research, (<http://www.driver-repository.eu/>)). Il progetto DRIVER è stato finanziato dal 6. e dal 7. Programma quadro e alla sua realizzazione partecipano 13 istituzioni che comprendono, oltre ai partner con ruolo tecnologico, i rappresentanti delle maggiori iniziative nazionali riguardanti l'*Open Access*.

Obiettivo di DRIVER è integrare in un'unica infrastruttura europea i contenuti degli archivi istituzionali e renderli accessibili in modo generalizzato attraverso servizi sofisticati ma, allo stesso tempo, di facile uso. E' infatti generalmente riconosciuto il ruolo che tali archivi possono giocare per la diffusione dei risultati della ricerca, ma la loro singola installazione risulta spesso troppo costosa, e con servizi di bassa qualità, talvolta oscurati da complicate procedure di accesso o di autorizzazione.

Il progetto dell'infrastruttura permette flessibilità e scalabilità, cosicché la costruzione dell'infrastruttura può avvenire partendo dalla realizzazione di singoli archivi, e procedere con la loro aggregazione in una struttura organizzata gerarchicamente secondo diversi criteri. Ripetendo questo processo in modo coerente si possono rendere accessibili con servizi avanzati tutti gli archivi europei e, potenzialmente, di tutto il mondo.

La prima fase si è conclusa con la realizzazione di una piattaforma sperimentale per la gestione di archivi di istituzioni scientifiche di Belgio, Francia, Germania, Olanda e Regno Unito. Seppur sperimentale, questa infrastruttura è stata progettata con criteri di alta qualità e si è basata su una architettura distribuita, orientata ai servizi. Essa ha permesso di ottenere un'alta interoperabilità nelle funzioni offerte a fornitori di servizi e contenuti. Il contenuto considerato in questa fase sperimentale è stato di tipo testuale: articoli, rapporti, preprints, etc.

Accanto a questa realizzazione, il Progetto ha fatto uno studio accurato della situazione esistente nei diversi paesi europei, rilevando le diverse sfaccettature che caratterizzano l'amministrazione degli archivi istituzionali, sia per quanto riguarda le tipologie dei documenti trattati, l'uso dei metadati, le politiche di accesso, le piattaforme tecnologiche utilizzate. Al fine di preparare il terreno per rendere interoperabili gli archivi distribuiti nell'infrastruttura, il Progetto ha definito linee guida rivolte sia alle organizzazioni che intendono esportare contenuti attraverso il protocollo OAI-PMH, sia agli amministratori degli archivi già esistenti, per sollecitarli allo sviluppo di servizi avanzati nell'ottica delle funzionalità dell'infrastruttura, sia agli sviluppatori di futuri archivi, per guidarli alla scelta di soluzioni tecnologiche che favoriscano l'interoperabilità.

Attualmente il contenuto di DRIVER consiste in 117 archivi istituzionali, ma il loro numero è destinato a crescere rapidamente.

La seconda fase vedrà la costruzione di una "production infrastructure", cioè una infrastruttura che garantisca servizi di qualità nell'arco delle 24 ore giornaliere, per 7 giorni la settimana e sarà caratterizzata da una particolare attenzione allo sviluppo di servizi orientati a comunità legate da interessi su specifici soggetti di ricerca. Gli sviluppi tecnologici riguarderanno essenzialmente la progettazione di funzionalità per il trattamento di oggetti digitali compositi, multimediali e multi-tipo, e di dati non strutturati.