



Open Access to Grey Resources

Nancy, France 5-6 December 2005 INIST-CNRS
Seventh International Conference on Grey Literature

Assisting scientists to make their research results
world wide freely available: an experience begun in
the 90s

Stefania Biagioni

Institute of Information Science and Technologies (ISTI)
Italian National Research Council (CNR)





TIFF (QuickTime™ and
are needed to see th

NCSTRL

Networked Computer Science Technical Reference Library

Contents

- The beginning
- The state of the art
- The evolution
- Conclusions





An experience begun in the 90s

- 1996 ERCIM* recognizes the importance of the DL technologies and decides to establish a Programme for R&D in DL sector
- The Italian CNR is appointed as coordinator of the “Ercim Digital Library Initiative” and is given the task of creating a research community in the Digital Library (DL) field in Europe within an international context
- One of its activities** is setting its own DL up to assist scientists to rapidly access, disseminate, manage and preserve over long term their intellectual output, in digital format, world wide over the web

*European Research Consortium for Informatics and Mathematics

**project funded by: ERCIM and DELOS Working Group on DL (ESPRIT Long Term Research Programme - LTR No. 21057)



An experience begun in the 90s (cont.)

1996-97 Ercim Librarians and Members of DL WG meet in Budapest and in Pisa and decide to build the the **ERCIM Technical Reference Digital Library**

as a specialized sub-collection of

NCSTRL - U.S. Networked Computer Science Technical Reference Library

with extended functionality to meet the requirements of a European Digital Library service

using **Dienst** technology (developed by US Cornell University DL Group) which allows:

networked access to a globally distributed collection of grey resources and
on-line services to assist scientists in publishing their work results, to review papers and administrate the repositories



An experience begun in the 90s (cont.)

The project, funded by ERCIM DELOS Working Group, was a collaborative effort among the following National Research Institutions:

CNR - Italy

CWI - The Netherlands

FORTH - Greece

GMD - Germany

INRIA - France

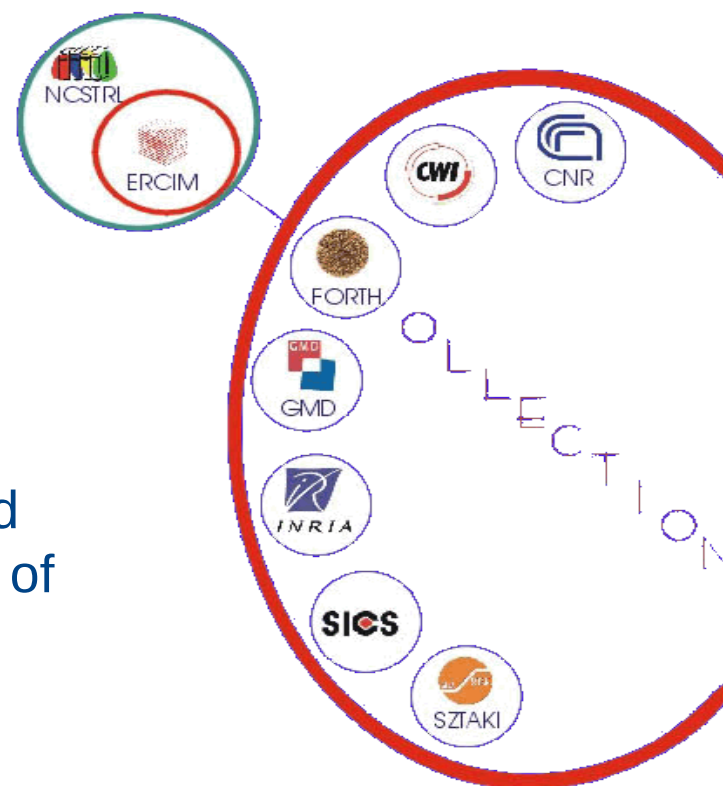
SICS - Sweden

SZTAKI - Hungary

ERCIM - Czech Republic

The system was developed by CNR and was operational after a one-year period of testing and refining by Ercim Librarians and developers

(1997-1999) (<http://dienst.isti.cnr.it>)



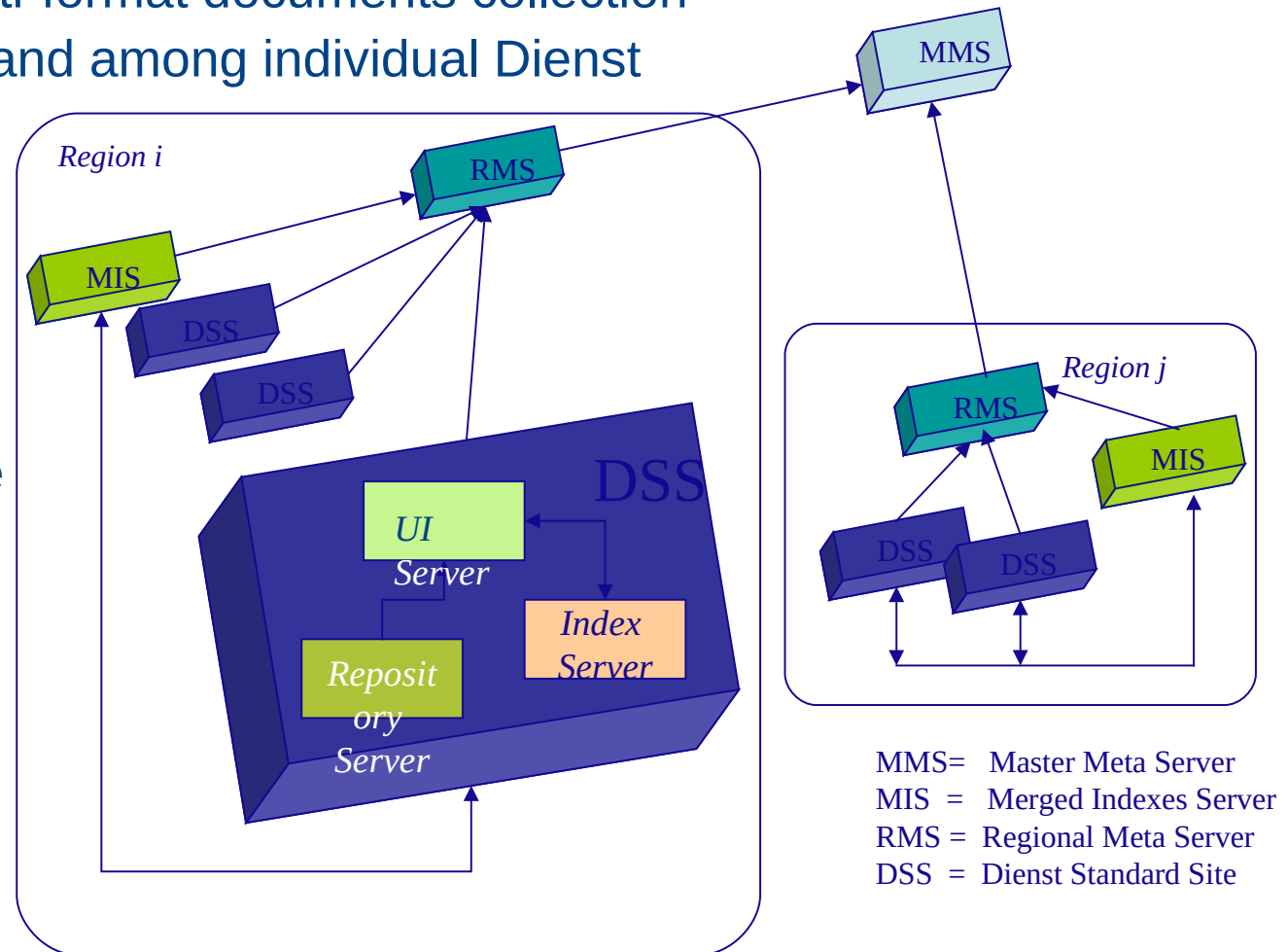
Dienst architecture for a federated digital library: NCSTRL

An open system that provides internet access to a distributed, decentralized, multi-format documents collection

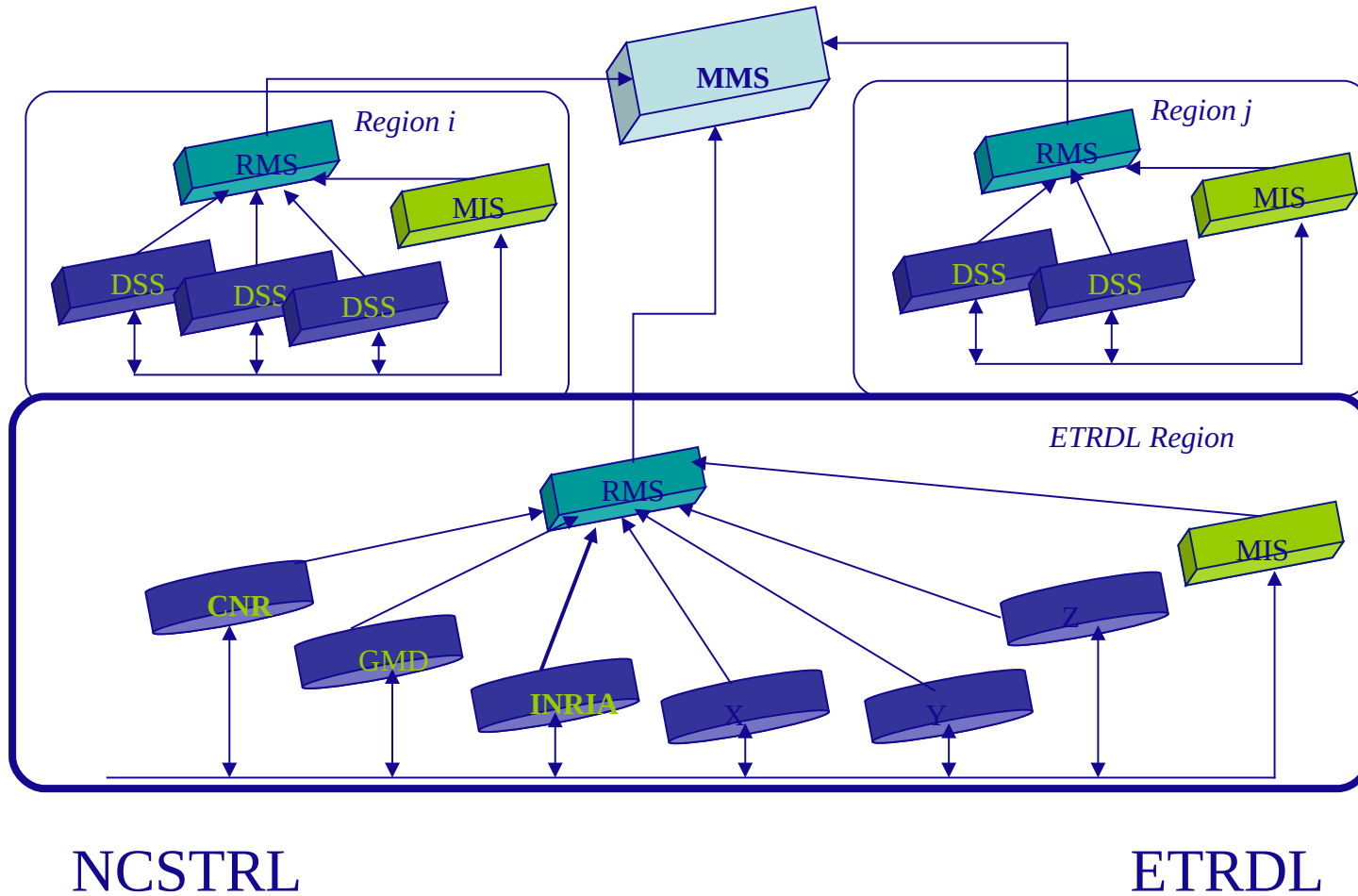
Communication with and among individual Dienst services takes place via an open protocol

Services defined in the protocol:

- Repository service
- Index service
- User interface Service
- Meta service



Dienst --> ETRDL region





A first experimental service

To provide an on-line service to assist Ercim scientists in exchanging their intellectual works and making it immediately world-wide

ETRD users

Information seekers, providers, administrators

ETRD documents (grey literature in CS and Math)

TR, EC deliverables, Pre-prints, Theses, Project reports, Ercim News, Proceedings ..

ETRD requirements

library service, not isolated, independent management of collections, language customization

ETRD access points

NCSTRL, ETRD, Local



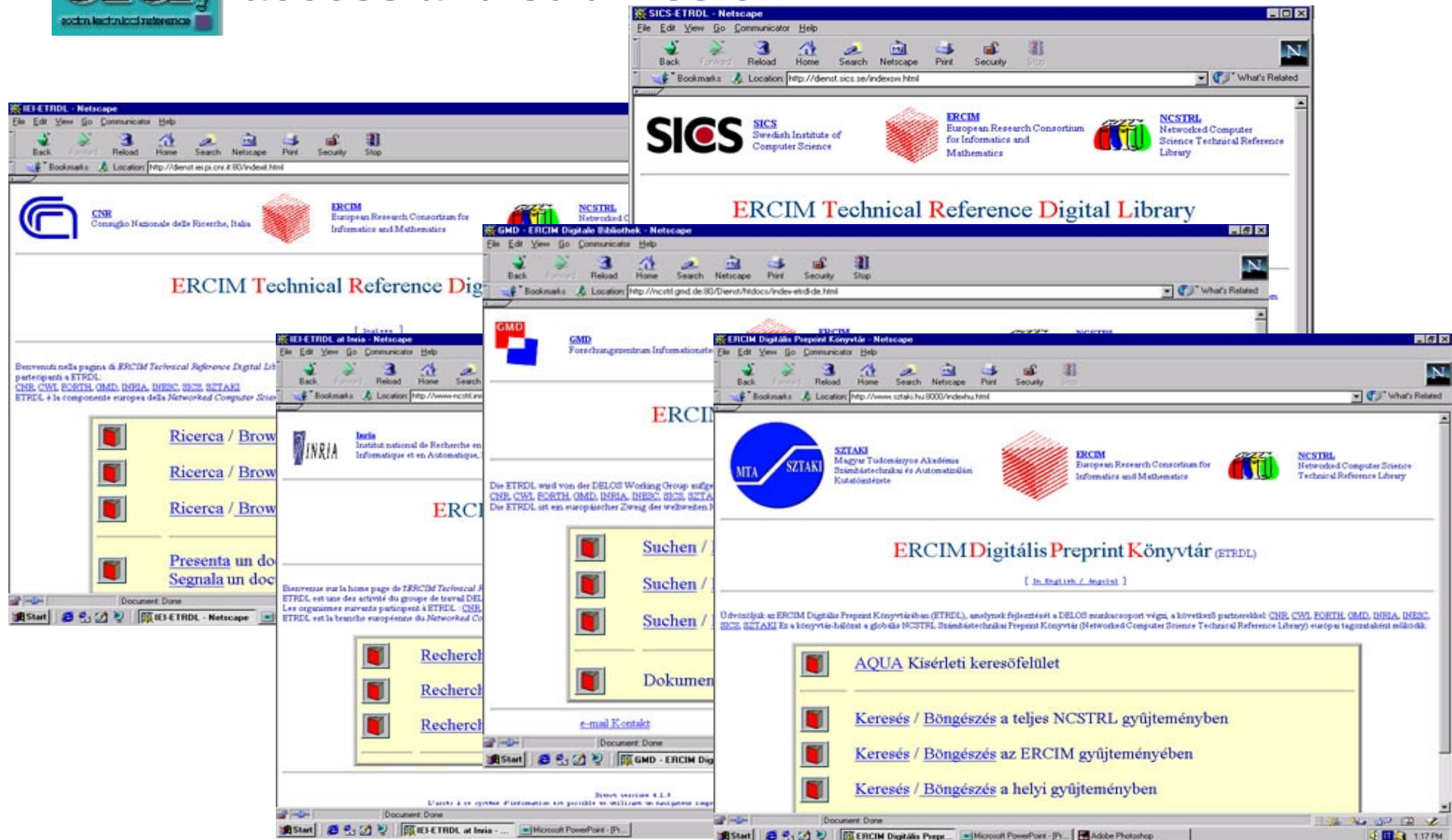
A specialized sub-collection of NCSTRL with additional functionality to meet the needs of a European DL

ETRDL extensions

- Multilingual information access
 - user interface localization - local language and English
 - different character set manipulation - Latin-1
 - cross language search
 - Metadata extensions (rfc 1807 vs DC)
- Search and Browse
 - Browsing extensions by subject (ACM, MSC, Free keywords)
 - Search extensions by subject, selection by date, language, type
 - more complex query structures
- Publishing - providing services for
 - submission, withdrawal
 - approval and administration



Pilot server with a multilingual information access and submission





Administration service

Partners are responsible for implementing their own administrative procedures

Here below we show the default system interface

The collage displays the ETRDL Administration Page interface across six overlapping screenshots. The top-left screenshot shows the login page with fields for Action, Collection, and Password. The top-middle screenshot shows the document management page with a list of documents and a 'View BIB Record' button. The top-right screenshot shows the document details page with a table containing fields like Title, Author(s), Date, Type, Language, Pages, Subject(s), and Keywords. The bottom-left screenshot shows the document reject page with a form for Title, Entry Date, Reason, and E-mail. The bottom-middle screenshot shows the document reject page with a form for Title, Entry Date, Reason, and E-mail. The bottom-right screenshot shows the document reject page with a form for Title, Entry Date, Reason, and E-mail.



Today: ETRDL is an operational service

- **ETRDL is**
 - a name of a software for specific applications
 - a self-publishing service for CNR Institutions
 - integrated in the context of library information systems
 - exchanges metadata with centralized CNR grey literature database
 - an access point to the new NCSTR

Together with **PUMA-DL**, a special instantiation collecting published works with controlled access, realizes the bibliography of some CNR Institutions

- **ETRDL-DEMO** has been set up for training courses



Today: special instantiations in the field of Marine Biology

- **MeCME DL - Mediterranean Costal Marine Environment** is a Digital Library made up in order to collect multidisciplinary documentation concerning scientific studies in the Mediterranean coastal areas and in particular in the Southern one, including also specific studies on Taranto Gulf. (IAMC-CNR-Taranto) <http://mecme.isti.cnr.it>
- **LVDL - Laguna di Venezia DL** is a multidisciplinary archive created in order to collect the multidisciplinary documentation arising from the scientific studies on the Lagoon of Venice. (University of Padua and Venice Universities, Naturalistic Observatory of the Lagoon and the Natural History Museum of Venice). <http://laguna.isti.cnr.it>

Further steps: web portal for the marine biology, migration to OAI-OpenDLib system, promotion of new collaborations in order to increase in number of collections



Evolution: migration to OpenDLib (<http://www.opendlib.com>)



- ETRDL needs a more advanced user oriented service in order to be compliant with currently requirements
- ETRDL will be managed by OpenDLib:
 - a new OAI compliant DL service system developed at ISTI following the experience with ETRDL

OpenDLib allows to create a digital library easily, according to the requirements of a given user community, by loading or harvesting the content to be managed



Evolution: migration to OpenDLib



OpenDLib is a federation of services which offers:

Main functionality of a digital library

submission, description, store, access, distribution, long term preservation,
peer-review, administration

Users management service

registration, maintaining of users profiles, authentication,
authorization

Infrastructure service

shared management and dynamic configuration of services



Evolution: migration to OpenDLib



OpenDLib is a federation of services which offers:

Personalization service

construction and dynamic modification of personalized virtual views of the information space

Virtual collections service

information space is organized in collections
collections are virtual and dynamically defined by authorized users
collections content is updated automatically
access rules and services may be associated to a collection



Evolution: migration to OpenDLib



OpenDLib can handle a variety of type of documents:
a powerful and flexible document model “DoMDL” capable to represent

multi-edition, structured, multimedia documents that can be disseminated in multiple manifestation formats

and allows any document to be associated

with multiple metadata descriptions in different formats and include semantic relationships with other documents and parts of them

OpenDLib architecture is designed to be interoperable with other DL
can act both as an OAI-PMH data and service provider



Evolution: migration to OpenDLib



First step in migrating ETRDL repositories in order to re-use metadata and documents

- mapping from RFC 1807 to Dublin Core format

Second step

- study of appropriate configuration of the system from architectural and semantic point of view:
 - which collections, which services
 - which additional metadata formats,
 - which document types with different formats
 - which search options: google-like or fielded, single or cross language, with or without relevance feedback

ETRDL implemented by a standard base configuration of OpenDLib: search service and document visualization

The screenshot displays the OpenDLib user interface in a Mozilla browser window. The main search results page shows a search for 'pagano' with a result set containing one entry. The entry details are as follows:

Field	Value
rank	73% ★★
handle	ercim.cnr.it/1997-ARCA-T14-TIM
version	1
dc:creator	Baldacci, Maria Bruna
dc:creator	Pagano, Pasquale
dc:date	May 4, 1999
dc:date	1997-06-02

A document viewer window is open, displaying the 'ARCA Target Configuration Interface Manual'. The document metadata is as follows:

Field	Value
Title	ARCA Target Configuration Interface Manual
Document Reference	ARCA/T14/TIM
Version	Version 1.0
Date	2 June 1997
Author(s)	Maria Bruna Baldacci Pasquale Pagano
Distribution	Intecs Sistemi CNR - IEI SABINI Fundacion Sancho El Sabio Universita' di Pisa Regione Toscana
Abstract	This is the manual for the users of the interface to configure the ARCA target upon a specific OPAC

The document viewer also shows a table of contents with the following entries:

Section	Page
body	1
Metadata information	1

The bottom of the screenshot shows the system tray with the date 'Wed Nov 16, 14:52' and the OpenDLib logo.

ETRDL new functionality: user management service

The screenshot displays the OpenDLib web interface in a Mozilla Firefox browser window. The main page, titled "ercim technical reference digital library", features a navigation menu on the left with links like Home, Community, Information space, Search, and Browse. The top navigation bar includes "My actions", "Users Management", "Groups", and "Docs Management".

The "Users rights management panel" is active, showing a table of users and their authorization status. The user "pagano (Pasquale Pagano)" is highlighted, indicating they are "Authorized to administrate this DL" (yes).

An inset window titled "http://metabib.isti.cnr.it:8000 - Edit user - Mozilla Firefox" provides a detailed view of "User rights for pagano". It includes radio buttons for "Authorized to administrate this DL" (set to "yes") and "Authorized to create collections" (set to "no"). Below this is a table of "Authorities" with their respective permissions.

Name	Authorized to administrate publishers	Authorized to approve submission to	Authorized to review documents	Authorized to request submission to
ercim.cnr.imc	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.iei	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.ilc	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.iroe	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.isti	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.iat	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.cnuce	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
ercim.cnr.iit	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

At the bottom of the inset window are "update" and "clear" buttons.

Conclusions

ETRDL powered by OpenDLib advantages

(in order to preserve metadata and documents against changes in technologies that could make them inaccessible and meaningless, allow the interoperability with other DL, handle multiple manifestation formats..)

- OAI compliancy
- Curation: automatic migration towards new formats and metadata
- Metadata and documents can be reused
- Archiving: preservation on an updated support, advanced according to current technologies
- Dynamic management of collections (also harvesting from other collections)
- Management of document relationships

Conclusions (cont.)

The experience made with ETRDL is a positive one

- From project phase to operational service integrated with library information services
- From original system to an advanced one
- Researchers became used to disseminate their work in a free way and associate different manifestations of documents to the same objects
- Repositories have been also used as test-bed for experimental activities

Conclusions (cont.)

Some difficulties have to be overcome yet at political, social, economic level

- Lack of coordination among different institutions
- Lack of knowledge of existing systems
- Lack of training among librarians
- Lack of being used to adopt such technology by scientists working in some fields of knowledge rather than others
- Lack of funds for involved activities

We need more consensus, interest and collaboration by academic, industrial and commercial organizations

Get in touch with ...

ETRDL / OpenDLib

Stefania Biagioni - ISTI CNR (stefania.biagioni@isti.cnr.it)

Carlo Carlesi - ISTI CNR (carlo.carlesi@isti.cnr.it)

Donatella Castelli - ISTI CNR (donatella.castelli@isti.cnr.it)

Pasquale Pagano - ISTI CNR (pasquale.pagano@isti.cnr.it)

Giuseppe Romano-ISTI CNR (giuseppe.romano@isti.cnr.it)

Manuele Simi - ISTI CNR (manuele.simi@isti.cnr.it)

LVDL

Antonella Zane - Vallisneri Library (antonella.zane@unipd.it)

Monica Ortolan - Vallisneri Library (monica.ortolan@unipd.it)

MeCME

Maria Filippi - Library of Talassografico Institute "A. Cerruti" IAMC CNR (maria.filippi@iamc.cnr.it)