

OpenDLib

Oltre le biblioteche digitali

Digital
Library
Service
System

Pasquale Pagano
ISTI-CNR
Pisa (Italy)

L'esperienza del CNR-ISTI

❖ ERCIM* WG on Digital Libraries
(1995-1996)

❖ ELOS WG on Digital Libraries
(1997-1999, 4th FP)

❖ ELOS NoE on Digital Libraries
(2000-2003, 5th FP)

❖ **Projects on Digital Libraries**
(2001-2003, 5th FP)

❖ **ELOS NoE on Digital Libraries**
(2004-2007, 6th FP)

❖ **Proposal on Digital Library & Grid**

* European Research Consortium for Informatics and Mathematics

Progetto EU

BRICKS - Building Resources for Integrated Cultural Knowledge Services

Progetto EU

DILIGENT - An Experimental Digital Library Infrastructure on Grid ENabled Technology

Diventare un punto di riferimento per il settore DL in Europa

Che cosa è una biblioteca digitale?

“ Una istituzione che fornisce (almeno) le funzioni di una biblioteca nel contesto di collezioni distribuite in rete di oggetti informativi in formato digitale ”

Nicholas Belkin
Eight DELOS Workshop
Stockholm, 1998

Obiettivo di una biblioteca digitale?

“ Le biblioteche digitali dovrebbero garantire ad ogni cittadino, indipendentemente dal luogo, la possibilità di accedere alla conoscenza in modo semplice, efficiente ed effettivo superando le barriere geografiche, linguistiche e culturali. ”

Brainstorming meeting
“Digital Library Research Directions”
organizzato da DELOS

OpenDLib: Funzionalità

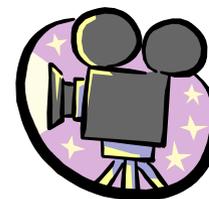
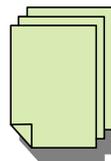
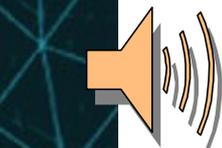
Un sistema che coordina servizi e risorse condivise in un ambiente dinamico e multi-istituzionale

- ❖ Acquisizione, memorizzazione, preservazione e disseminazione
- ❖ Autorizzazione e autenticazione
- ❖ Analisi, ricerca, ed accesso

- ❖ Organizzazione dello spazio informativo
- ❖ Valorizzazione
- ❖ Interoperabilità

Oggetti Digitali

Ogni documento, o insieme di documenti, riconducibile ad un'istituzione che abbia valore ai fini della conservazione e del recupero dell'informazione



Oggetti Digitali

Un oggetto digitale è un **contenitore** che contiene altri oggetti digitali correlati

Ad ogni contenitore e ad ogni suo parte è possibile associare una **descrizione** che ne evidenzia aspetti peculiari

Le descrizioni gestibili, **metadati**, sono strutturate, o semistrutturate, e permettono la gestione, il recupero e l'uso degli oggetti digitali e delle loro parti

Oggetti Digitali

L'insieme dei documenti prodotti o ricevuti da una persona/organizzazione durante la propria attività

Gli oggetti vengono gestiti tenendo conto del:

❖ **Tipologia del contenuto**

❖ **Struttura**

- Insieme di relazioni che collegano i diversi documenti secondo un percorso ben definito

❖ **Contesto**

- Insieme delle informazioni relative alla “storia” dei documenti

❖ **Presentazione**

- Insieme di vincoli che regolano l'accesso all'oggetto tenendo conto del contenuto, della struttura e del contesto nonché dell'utente che ha effettuato la richiesta

Funzionalità: acquisizione

In tale fase i documenti possono essere:

- ❖ Classificati in accordo alla schema di classificazione scelto dall'Istituzione
- ❖ Correlati ad altri **Oggetto Digitale** esistente
- ❖ Descritti, a livello di insieme e di documento, in accordo ai formati di metadati adottati dall'Istituzione

Funzionalità: mantenimento

I documenti acquisiti **non** possono essere modificati. Le correzioni, modifiche o aggiunte di informazioni ad un documento, se autorizzate, vengono gestite attraverso le versioni

Gli oggetti digitali possono essere:

- ❖ Duplicati automaticamente in base alle specifiche istituzionali, al fine di garantire l'integrità referenziale
- ❖ Cancellati da un amministratore nell'ambito di una procedura controllata
- ❖ Distrutti/Archiviati automaticamente in base al programma di conservazione

Funzionalità: preservazione

OpenDLib gestisce il fenomeno della **obsolescenza tecnologica** mediante

1. La conversione automatica dei dati e dei metadati verso formati standard
2. La migrazione automatica dei dati e dei metadati verso formati sempre aggiornati
3. La migrazione del software necessario per accedere, trasmettere e visualizzare le informazioni

Funzionalità: autorizzazione

Le Istituzioni, attraverso un loro amministratore, assegnano agli utenti la possibilità di

❖ accesso

- a determinate collezioni
- agli oggetti digitali
- alle parti di un oggetto
- ai metadati o loro parte

❖ utilizzazione di funzioni particolari

- lettura
- aggiornamento
- cancellazione
- inserimento

Funzionalità: accesso

I documenti acquisiti possono essere fruiti da tutti gli utenti che ne hanno diritti

Il servizio di Accesso modula l'accesso tenendo conto:

- ❖ del profilo dell'utente
- ❖ del contesto in cui viene effettuata la richiesta
- ❖ delle regole di accesso poste
 - sull'oggetto digitale
 - su una sua parte
 - sui singoli documenti collegati
 - sui metadati o su parte di essi
 -

Funzionalità: ricerca

Sono ricercabili tutte le informazioni associate agli oggetti e alle loro parti:

- ❖ Rappresentazioni del contenuto
- ❖ Riferimenti ad altri oggetti
- ❖ Contenuto di oggetti digitali in formato testo

Le operazioni di ricerca e gli elementi ricercabili sono personalizzabili per categorie di utenti

Tipologia delle operazioni di ricerca supportate

1. Per parole chiave
2. Strutturata



The screenshot shows a search interface with a table of search options and a search form below it.

Result set	Search options	Search type
		Simple
		Advanced

L'utente può raffinare i risultati delle ricerche sia con ulteriori operazioni di ricerca, sia avvalendosi della funzionalità di "relevance-feedback"

search terms:

search clear

Spazio Informativo

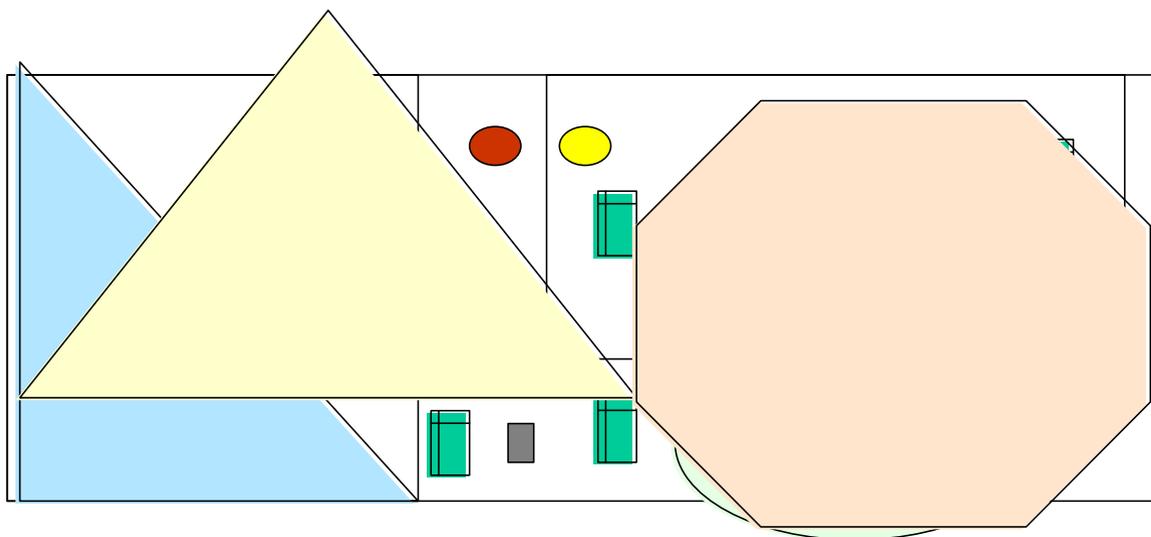
Lo spazio informativo di una biblioteca digitale è costituito dall'insieme delle risorse accessibili.

Tale spazio riferisce normalmente oggetti eterogenei sia per struttura, formato, contenuto semantico, diritti di accesso, che per la potenziale audience a cui sono rivolti.

Lo spazio informativo non può essere definito staticamente ma deve essere modellabile e personalizzabile dall'amministratore della biblioteca digitale che ne stabilisce la struttura e l'organizzazione.

Spazio Informativo : **collezioni**

Le collezioni permettono di personalizzare l'organizzazione degli oggetti digitali in modo da soddisfare le esigenze permanenti delle diverse comunità di utenti



[cont]

Analizzare

Manage	View	Personal
	Public collections	
	Personal collections	
	Current information space	

Personalizzare la sessione di lavoro

Manage	View	Personal
Select ->	from public collections	
Remove	from personal collections	

Personalizzare lo spazio informativo

Manage	View	Personal
		Create collection
		Delete collection

Valorizzazione

- ❖ Identificazione dei metadati nativi necessari per supportare la ricerca
- ❖ Definizione delle regole di riscrittura dei due formati in Dublin Core per garantire l'interoperabilità
- ❖ Progettazione del modello del documento per garantire una effettiva ed efficace navigazione della conoscenza
- ❖ Identificazione e supporto delle relazioni tra gli oggetti

Valorizzazione



[cont.]

Fonti iconografiche

Metadata

Immagine

Relazioni

Fonte testuale

Fonti iconografiche
simili

Handle	sns.it/2004_F_171
Version	1
<input type="checkbox"/> metadata	
dc	view download
memoria	view download
<input type="checkbox"/> immagine	
illustrazione	view download
<input type="checkbox"/> relazioni	
<input type="checkbox"/> contenuto_in	
<input type="checkbox"/> correlato_con	
<input type="checkbox"/> MargaritaPhilosophic-pag625	
<input type="checkbox"/> NumerazioneDeiLuoghi1	
<input type="checkbox"/> SimboliPlanetariEAst2	
<input type="checkbox"/> metadata	
<input type="checkbox"/> immagine	
<input type="checkbox"/> MostroMnemonic3	

[cont.]

Fonti testuali

Handle	sns.it/2004_F_80
Version	1
<input type="checkbox"/> metadata	
dc	view download
memoria	view download
<input type="checkbox"/> relazioni	
<input type="checkbox"/> contiene	
<input type="checkbox"/> RebusMnemonico-pag18	
<input type="checkbox"/> AlfabetoMnemonicoFor-pag19	
<input type="checkbox"/> AlfabetoMnemonicoPos-pag18	
<input type="checkbox"/> metadata	
<input type="checkbox"/> immagine	
illustrazione	view download
<input type="checkbox"/> ImmaginiDeiNumeri-pag20	
<input type="checkbox"/> RebusMnemonico-pag17	
<input type="checkbox"/> riferito_da	
<input type="checkbox"/> Arskeminiscendi1	
<input type="checkbox"/> metadata	
<input type="checkbox"/> relazioni	

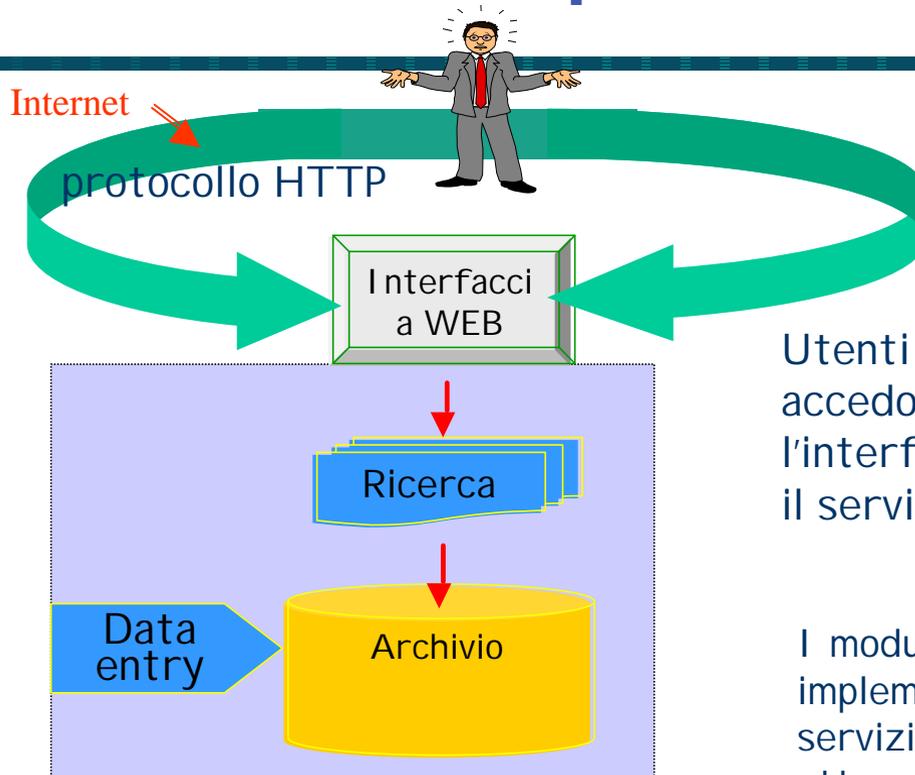
Metadata

Relazioni

Fonti iconografiche
contenute

Fonti testuali
correlate

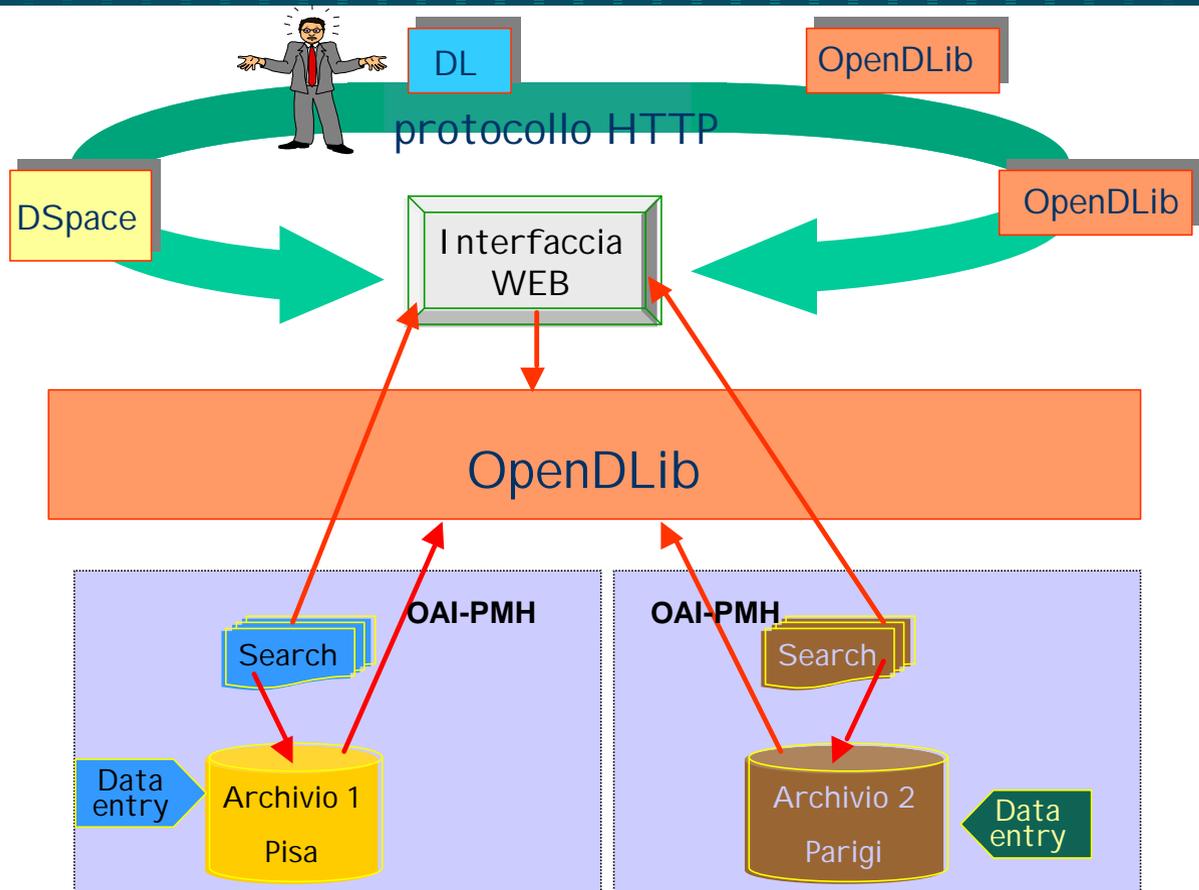
Interoperabilità



Utenti generici
accedono attraverso
l'interfaccia Web solo
il servizio di search

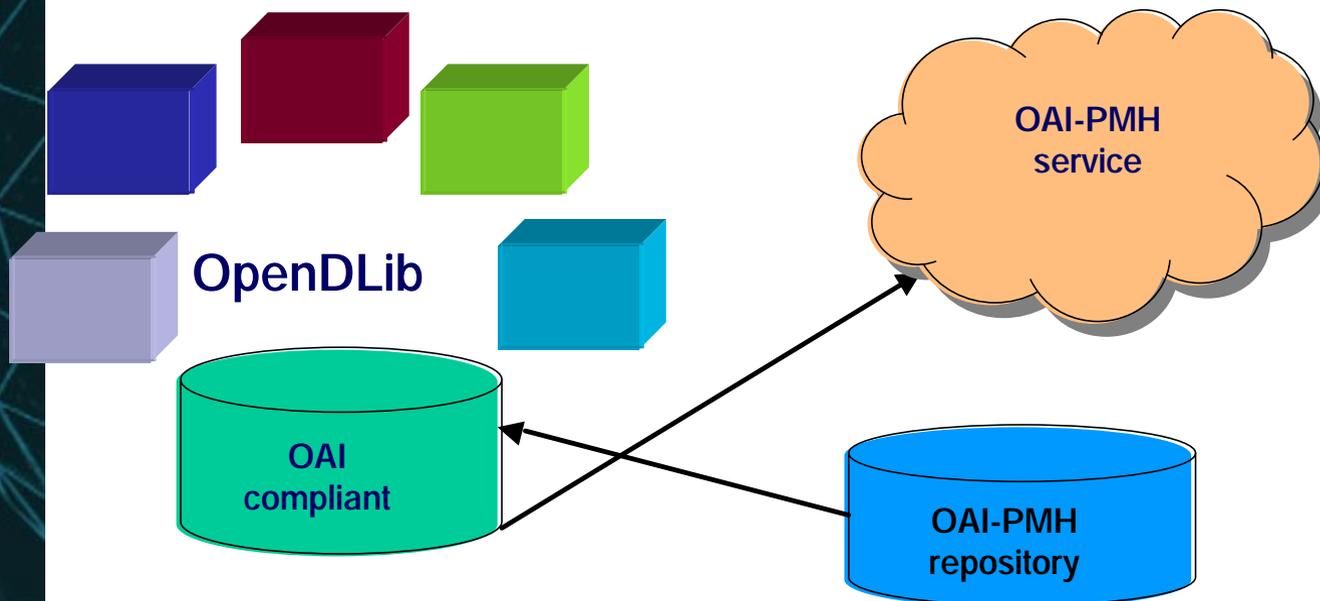
I moduli software che
implementano gli altri
servizi non comunicano
attraverso il Web

Interoperabilità



Interoperabilità

- ❖ Una biblioteca OpenDLib può recuperare metadati da altre fonti utilizzando il protocollo OAI-PMH
- ❖ I metadati mantenuti in una biblioteca OpenDLib possono essere resi accessibili tramite il protocollo OAI-PMH



Ulteriori informazioni

Biblioteca digitale di dimostrazione disponibile
all'indirizzo specificato in

<http://www.opendlib.com>

Contact people:

- ❖ Donatella Castelli (donatella.castelli@isti.cnr.it)
- ❖ Pasquale Pagano (pasquale.pagano@isti.cnr.it)

1995 - Budapest: **primo incontro fra Bibliotecari ERCIM e membri del Gruppo Digital Libraries**

- vengono gettate le basi per la realizzazione di una Biblioteca Digitale, in collaborazione con la Cornell University (NCSTRL*), finalizzata alla diffusione della letteratura scientifica del settore Information Technology

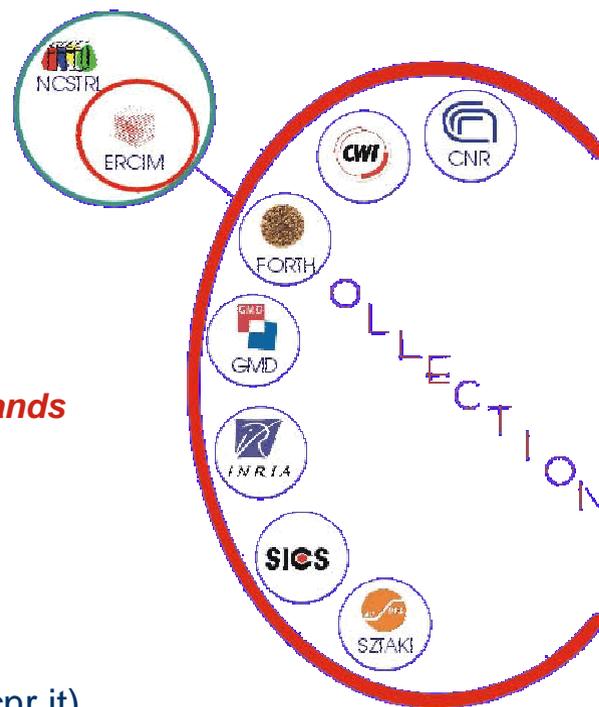
1996 - ISTI-CNR diventa coordinatore della “**ERCIM Digital Library Initiative**” con il compito di creare una comunità di ricerca nel campo delle Digital Library (DL) in Europa

Networked Computer Science Technical Reference Library

DELOS WG on Digital Libraries (4th FP)

1997-1999 Ercim Technical Reference Digital Library

Sotto-insieme di NCSTRL
con funzionalità estese
per soddisfare le
necessità dell'ambiente di
ricerca europeo
(interfacce multilingua)



CNR - Italy
CWI - The Netherlands
FORTH - Greece
GMD - Germany
INRIA - France
SICS - Sweden
SZTAKI - Hungary



(<http://etrdl.iei.pi.cnr.it>)

- ❖ Ha un'architettura distribuita e multi formato
- ❖ Realizza un modello alternativo di editoria scientifica
- ❖ Offre i servizi tradizionali di una biblioteca su *documenti tradizionali in formato digitale*
 - permette una disseminazione più rapida, più ampia, meno costosa
- ❖ E' uno strumento per la creazione di collezioni di documenti di qualità controllata
 - immediatamente accessibili in rete
 - gestite dagli autori stessi, e dalle loro istituzioni, nella veste di produttori ed editori delle informazioni

- ❖ Attualmente e' un servizio attivo nell'area Informatica e Matematica
- ❖ Sono state, inoltre, attivate due istanze personalizzate per aree disciplinari diverse
 - **LVDL - Laguna di Venezia DL** (Biologia Marina) : DL tematica finalizzata a raccogliere la documentazione multidisciplinare relativa a studi dell'ambiente lagunare veneziano
 - **HF-EPR Spectroscopy DL** (Fisica) : DL tematica finalizzata a raccogliere la documentazione dei partecipanti al progetto Infrastructure Cooperation Network SENTINEL / EU coordinato dall'IPCF-CNR

Progetti per la diffusione della conoscenza su Digital Libraries

2001-2003

- ❖ OPEN ARCHIVE FORUM - un forum per la disseminazione dell'informazione sulle attività eropee collegate alla Open Archive Initiative
- ❖ DLib Competence Center on DL - fornisce a specifiche comunità l'accesso a sistemi avanzati per DL, tecnologie, servizi, esperienza e conoscenza

Progetti di ricerca e sviluppo su Digital Libraries

2001 - 2003

- ❖ ECHO - European CHronicles On-line
 - un sistema di Biblioteca Digitale di documentari storici audio-video basato su tecniche di indicizzazione automatica

- ❖ CYCLADES - An Open Collaborative Digital Library Service Environment accessing Open Archives Initiative Compliant Information Sources
 - un sistema per lo scambio di conoscenza, ricerche avanzate, realizzazione di collezioni virtuali, profilazione e raccomandazione

Progetti di ricerca e sviluppo su Digital Libraries (cont.)

- ❖ SCHOLNET - A Digital Library Testbed to Support Networked Scholarly Communities
 - un sistema per
 - la gestione di tipi dati non solo testuali
 - l'annotazione ipermediale
 - la ricerca cross-language
 - la disseminazione personalizzata dell'informazione
 - un sistema basato su un software toolkit, [OpenDLib](#), che supporta DL federate i cui moduli realizzano servizi avanzati e personalizzabili

2004 - 2007 Cooperazione con l'ambiente DL nazionale e internazionale

❖ Nuovi progetti EU:

● **DELOS NoE on Digital Libraries**

■ Dlib Competence Center

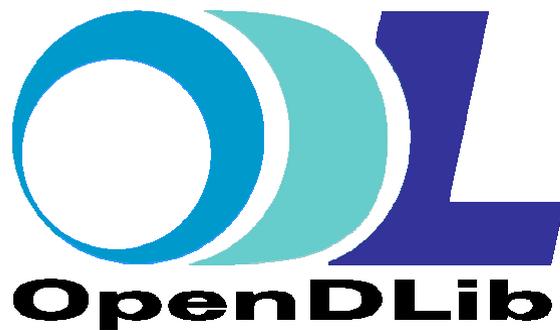
● **BRICKS - Building Resources for Integrated Cultural Knowledge Services**

❖ Nuove proposte EU:

● **DILIGENT - An Experimental Digital Library Infrastructure on Grid Enabled Technology**

❖ Nuovi Servizi:

● **Attivazione di OpenDLib:** il nuovo software DL toolkit sviluppato presso l'ISTI



<http://www.opendlib.com>

OpenDLib - Motivazioni

- ❖ Le biblioteche digitali sempre di più vengono percepite come uno strumento per supportare la comunicazione e la co-operazione tra comunità di ricercatori

- ❖ Queste comunità
 - sono relativamente piccole ed hanno risorse limitate
 - non hanno esperienza informatica
 - hanno bisogno in tempi brevi di una biblioteca digitale operante



OpenDLib – Motivazioni (cont.)

OpenDLib è stato progettato per permettere di soddisfare queste richieste supportando un modello di sviluppo e mantenimento di biblioteca digitale “a basso costo”

OpenDLib è un toolkit che può essere usato per creare biblioteche digitali

- ❖ il software viene configurato secondo le necessità della comunità di utenti
- ❖ i documenti digitali e i relativi record bibliografici possono essere sottomessi direttamente oppure importati da fonti esistenti

OpenDLib - Architettura

OpenDLib è una federazione di servizi che offrono:

- ❖ funzionalità di base di una biblioteca digitale
 - sottomissione, descrizione, ricerca, browsing, recupero, accesso, preservazione, interfaccia utente
- ❖ servizi per la gestione degli utenti
 - registrazione, mantenimento dei profili utente, autenticazione, autorizzazione
- ❖ servizi per la personalizzazione
 - costruzione e modifica dinamica di viste virtuali personalizzate dello spazio informativo e dei servizi su di esso
- ❖ servizi infrastrutturali

La federazione è open, può essere estesa aggiungendo altri servizi specifici richiesti dalla comunità di utenti

Gestione condivisa

- ❖ La federazione può comprendere istanze multiple di un stesso servizio, anche configurate diversamente
- ❖ La federazione può essere gestita da una unica organizzazione oppure da un consorzio di organizzazioni
- ❖ Ogni partecipante al consorzio può decidere quali istanze di servizio mantenere localmente sui propri computer

Configurazione dinamica

- ❖ La distribuzione e il numero delle istanze di uno stesso servizio, la loro configurazione e allocazione possono essere cambiati dinamicamente per far fronte a nuove esigenze

Esempi:

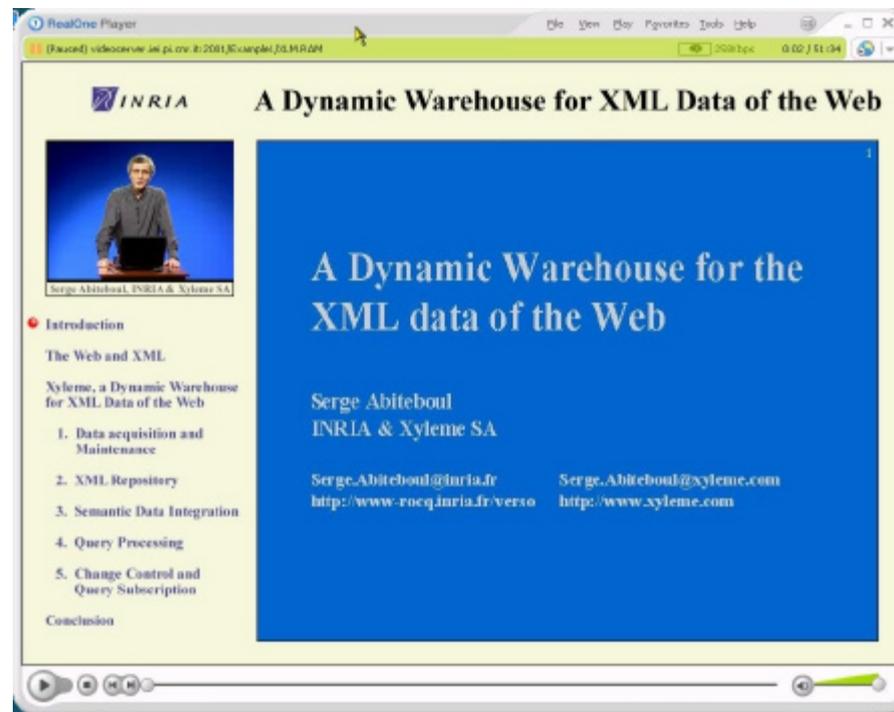
- Una nuova istanza del servizio di indicizzazione può essere aggiunta per ridurre il carico su quelle esistenti e migliorare i tempi di risposta
- Un nuovo servizio di memorizzazione può essere aggiunto per memorizzare i documenti prodotti da una nuova organizzazione che si aggiunge al consorzio
- Un nuovo servizio di gestione delle interrogazioni può essere aggiunto per supportare lo specifico linguaggio di interrogazione di una nuova comunità di utenti

Tipo dei documenti (1)

- ❖ Una biblioteca OpenDLib può mantenere documenti che differiscono per formato, struttura e media. La stessa biblioteca può mantenere, ad esempio:
 - Una collezione di riviste e atti di conferenze composte da articoli
 - Una collezione di tesi organizzate in capitoli e sezioni
 - Una collezione di video strutturati in sequenze e shot
 - Una collezione di altro materiale rappresentata dalle corrispondenti registrazioni catalografiche

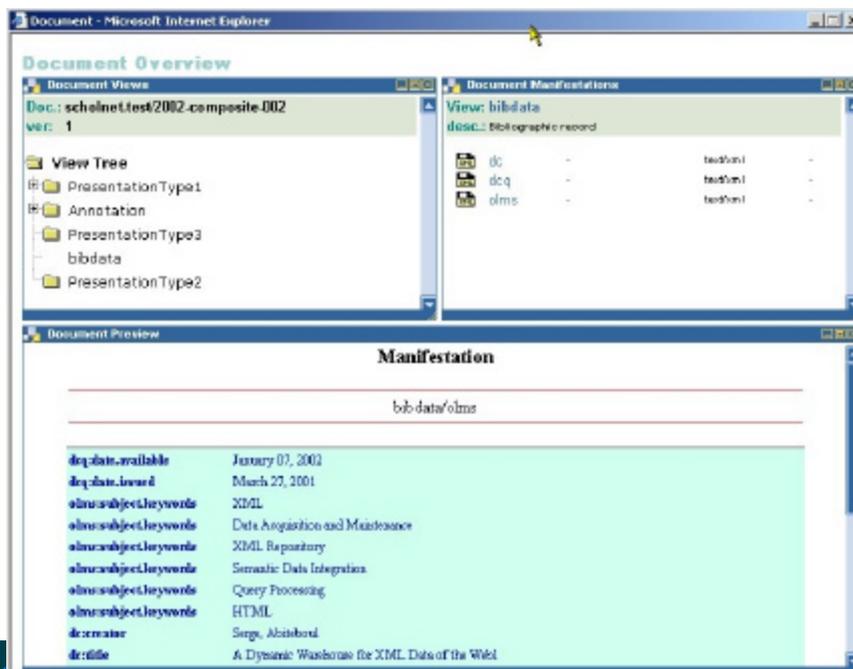
Tipo dei documenti (2)

- ❖ OpenDLib può anche gestire nuovi tipi di documenti che non hanno una controparte fisica



Tipo dei documenti (3)

- ❖ OpenDLib può mantenere edizioni, versioni e manifestazioni multiple di uno stesso documento
- ❖ Più descrizioni catalografiche in formati diversi possono essere associate ad uno stesso documento



The screenshot shows a web browser window titled "Document - Microsoft Internet Explorer" displaying a "Document Overview" page. The page is divided into three main sections:

- Document Views:** Shows document details: "Doc.: scholnet/test2002-composite-002" and "ver.: 1".
- View Tree:** A hierarchical tree structure with folders: "PresentationType1", "Annotation", "PresentationType3", "bibdata", and "PresentationType2".
- Document Manifestations:** A table showing different views of the document. The selected view is "bibdata" with destination "Bibliographic record".

The "Document Manifestations" table contains the following data:

View	Format	Manifestation	Thumbnail
dc	-	test001	-
dcq	-	test001	-
olms	-	test001	-

The "Document Preview" section shows a "Manifestation" titled "bib data/olms". Below the title is a table of metadata:

dc:date.available	January 07, 2002
dc:date.issued	March 27, 2001
olms:subject.keywords	XML
olms:subject.keywords	Data Acquisition and Maintenance
olms:subject.keywords	XML Repository
olms:subject.keywords	Semantic Data Integration
olms:subject.keywords	Query Processing
olms:subject.keywords	HTML
dc:creator	Serge, Abitchev
dc:title	A Dynamic Warehouse for XML Data of the Web

Collezioni virtuali

- ❖ Lo spazio dei documenti è organizzato in collezioni
- ❖ Le collezioni sono virtuali e possono essere definite dinamicamente da chiunque è autorizzato a farlo
 - Rapporti tecnici prodotti dagli istituti del CNR
 - Tesi di dottorato in inglese
 - Immagini prodotte dai membri del gruppo di lavoro su biologia marina
- ❖ Il contenuto di una collezione viene aggiornato automaticamente
- ❖ Posso associare ad una collezione regole di accesso e servizi – la collezione è di fatto una sorta di biblioteca virtuale