The DL.org Reference Model for Digital Libraries

Vittore Casarosa
ISTI-CNR, Pisa, Italy

“Digital Library and Open Access. Interoperability strategies” Workshop
London, 4 February 2011
DL.Org and DELOS
The Network of Excellence on Digital Libraries

Define and conduct a **joint program of activities** in order to integrate and coordinate the on-going research activities of the major European research teams in the field of **digital libraries** for the purpose of developing the next generation digital library technologies.
Digital libraries should enable any citizen to access all human knowledge anytime and anywhere, in a friendly, multi-modal, efficient, and effective way, by overcoming barriers of distance, language, and culture and by using multiple Internet-connected devices.
The potential exists for digital libraries to become the universal knowledge repositories and communication conduits for the future, a common vehicle by which everyone will access, discuss, evaluate, and enhance information of all forms.
Digital Libraries in the Information Space - 1

Structure of Data

Structure of User Behavior

Low

High

Databases/IR

Digital Libraries

Web

High
Structure of Data

- Info Retrieval
- Databases
- CMS/DAMS
- Wikis/blogs
- Digital Libraries
- Web

Structure of User Behavior

High

Low

Digital Libraries in the Information Space - 2

London
February 4th, 2011

DL.org Workshop - Vittore Casarosa
Digital Libraries and Open Access. Interoperability strategies
Conceptual Framework

Digital Library System

Contents

Management

Usage
Research Directions in Digital Libraries

- Applications: e-Health, e-Learning, e-Culture, etc.
- User:
  - User Interfaces
  - Visualization
  - Personalization
- Systems:
  - Architectures
  - Information Access
  - Audio/Visual
  - Semantic Interop
- Foundations:
  - Reference Model for DLS
What is a Digital Library?

- A DL is the combination of “content and services”
- A Digital Library is a “tool” at the centre of intellectual activity having no logical, conceptual, physical, temporal, or personal borders or barriers on information
- A DL is an “entity” providing the functionality to mediate between information objects and information users in the context of distributed collections of information objects. This (external) functionality includes access, publish, delivery, preservation, personalization, etc.
- A Digital Library is an “institution” in charge of providing at least the functionality of a traditional library in the context of distributed and networked collections of information objects
Need for a Reference Model

• A reference model is an **abstract framework** for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment

• A reference model is based on a **small number of unifying concepts** and may be used as a basis for education and explaining standards to a non-specialist

• A reference model **is not directly tied to any standards, technologies or other concrete implementation details**, but it does seek to provide a common semantics that can be used unambiguously across and between different implementations
A Three-Entity Framework
The three Entities

• Digital Library
  – An organization, which might be virtual, that comprehensively collects, manages, and preserves for the long term rich digital content, and offers to its user communities specialized functionality on that content, of measurable quality and according to codified policies

• Digital Library System
  – A software system that is based on a defined (possibly distributed) architecture and provides all functionality required by a particular Digital Library. Users interact with a Digital Library through the corresponding Digital Library System

• Digital Library Management System
  – A generic software system that provides the appropriate software infrastructure both (i) to produce and administer a Digital Library System incorporating the suite of functionality considered foundational for Digital Libraries and (ii) to integrate additional software offering more refined, specialized, or advanced functionality
Different types of DLMS

• **Extensible Digital Library System**
  – A complete Digital Library System that is fully operational with respect to basic/foundational functionality required. It is based on an open software architecture, so that further software components can be incorporated on top of the ones already there with ease (DelosDLMS, GreenStone)

• **Digital Library System Warehouse**
  – A collection of software components that encapsulate the core suite of DL functionality and a set of tools that can be used to combine these components in a variety of ways (in Lego-like fashion) to create Digital Library Systems offering a tailored integration of functionalities. New software components can easily be incorporated into the Warehouse for subsequent combination with those already there (BRICKS, DILIGENT)

• **Digital Library System Generator**
  – A highly parameterized software system that encapsulates templates covering a broad range of functionalities, including a defined core suite of DL functionality as well as any advanced functionality that has been deemed appropriate to meet the needs of the specific application domain. Through an initialization session, the appropriate parameters are set and configured; at the end of that session, an application is automatically generated, and this constitutes the Digital Library System ready for installation and deployment (MARIAN)
DELOS DLMS

OSIRIS – Service-oriented DL Framework (Integration Platform)

Documents and Collections

Multilinguality

Visualization

Ontology

Audio Retrieval

Index

Video Retrieval

Personalization

ISIS Services

DL.org Workshop - Vittore Casarosa
Digital Libraries and Open Access. Interoperability strategies
Actors in the Digital Library

Digital Library

Digital Library System

Digital Library Management System

DL End-Users

DL Designers

DL System Administrators

DL Application Developers
Main roles of Actors (1)

• **DL End-Users**
  - They exploit the DL functionality for providing, consuming, and managing the DL Content as well as some of its other constituents. They perceive the DL as a stateful entity that serves their functional needs. The behaviour and output of the DL depend on its state at the time a particular part of its functionality is activated. DL end-users may be further partitioned into
    • Content Creators
    • Content Consumers
    • Librarians (end user)

• **DL Designers (Digital Librarian)**
  - They exploit their knowledge of the semantic of the application domain to define, customize, and maintain the Digital Library so that it is aligned with the information and functional needs of its end-users. To perform this task, they interact with the DLMS providing functional and content configuration parameters. The values of these parameters, which can be modified during the DL lifetime, configure the specific DL perceived by the end-users because they determine the particular Digital Library System instance serving the Digital Library.
Main roles of Actors (2)

- **DL System Administrators (System Librarian)**
  - They select the software components necessary to create the Digital Library System needed to serve the required DL (as specified by the DL Designer) and decide where and how to deploy them. They interact with the DLMS by providing architectural configuration parameters, such as the selected software components, the hosting nodes, and the components allocation. The value of the architectural configuration parameters can be changed over the DL lifetime. Any change of these parameters may result in the provision of different DL functionality and/or different quality.

- **DL Application Developers**
  - They develop the software components of DLMSs and DLSs, implementing the necessary functionality.
Main concepts (1)
Main concepts (2)

• **Content**
  – The Content concept encompasses the data and information that the Digital Library handles and makes available to its users. Content is an umbrella concept used to aggregate all forms of information objects that a Digital Library collects, manages, and delivers. It encompasses the diverse range of information objects, including such resources as objects, annotations, and metadata.

• **User**
  – The User concept covers the various actors (whether human or machine) entitled to interact with Digital Libraries. Digital Libraries connect actors with information and support them in their ability to consume and make creative use of it to generate new information. User is an umbrella concept including all notions related to the representation and management of actor entities within a Digital Library. It encompasses such elements as the rights that actors have within the system and the profiles of the actors with characteristics that personalize the system’s behaviour or represent these actors in collaborations.

• **Functionality**
  – The Functionality concept encapsulates the services that a Digital Library offers to its different users, whether classes of users or individual users. While the general expectation is that DLs will be rich in capabilities and services, the bare minimum of functions would include such aspects as new information object registration, search, and browse. Beyond that, the system seeks to manage the functions of the Digital Library to ensure that the functions reflect the particular needs of the digital library’s community of users and/or the specific requirements relating to the Content it contains.
Main concepts (3)

- **Policy**
  - The Policy concept represents the set (or sets) of conditions, rules, terms and regulations governing interaction between the Digital Library and users, whether virtual or real. Examples of policies include acceptable user behaviour, digital rights management, privacy and confidentiality, charges to users, and collection delivery.

- **Quality**
  - The Quality concept represents the parameters that can be used to characterize and evaluate the content and behaviour of a Digital Library. Quality can be associated not only with each class of content or functionality but also with specific information objects or services. Some of these parameters are objective in nature and can be automatically measured, whereas others are subjective in nature and can only be measured through user evaluations.

- **Architecture**
  - The Architecture concept refers to the Digital Library System entity and represents a mapping of the functionality and content offered by a Digital Library onto hardware and software components. There are two primary reasons for having Architecture as a core concept: (i) Digital Libraries are often assumed to be among the most complex and advanced forms of information systems; and (ii) interoperability across Digital Libraries is recognized as a substantial research challenge. A clear architectural framework for the Digital Library System offers ammunition in addressing both these issues effectively.
The main concepts in perspective
The Digital Library Universe

3 Types of systems
• DL
• DL System
• DL Management Systems

6 Domains + 1
• Content
• User
• Functionality
• Policy
• Quality
• Architecture
• + Resource

4 Roles of Actors
• DL end-Users
• DL Application Developers
• DL Designers
• DL Systems Administrators
Concept Maps

- Concept Maps help to answer needed for effective teaching and learning.
- Organized Knowledge includes concepts, linked words, propositions, and more.
- Context dependent on effective teaching and learning, e.g., personal and social.
- Interrelationships between different map segments.
- Creativity aids especially with cognitive structure and experts.
- Infants begin with creativity and symbols.
- Events (Happenings) and objects (Things) are connected with perceived regularities or patterns.
Digital Library Domains

218 concepts
52 relationships
The Resource Domain

[Diagram showing relationships between Digital Library, Resource Domain, Purpose, Resource, Resource Set, Information Object, and Ontology with various properties like definedBy, consistsOf, associatedWith, hasPart,belongsTo, isa, identifiedBy, hasFormat, hasQuality, regulatedBy, hasAnnotation, hasMetadata, describedBy, expressedBy, expressionOf, expressionOf, Policy, Quality Parameter, Region, and Ontology.]
The Content Domain
The User Domain
The Functionality Domain

Digital Library

Digital Library System

Digital Library Management System

support

deploy

extend

offer

actOn

interactWith

purpose

associatedWith

hasPart

resource

identifiedBy

regulatedBy

hasQuality

hasAnnotation

hasMetadata

Region

Policy

Quality Parameter

Resource

Information Object

Actor

modelledBy

Actor Profile

Manage Resource

Collaborate

Manage DL

Manage & Configure DLS

influencedBy

perform

isa
Main functions

C32 Access Resource
  C33 Discover
    C34 Browse
    C35 Search
  C36 Acquire
  C37 Visualize

C38 Manage Resource
  C39 Create
  C40 Submit
  C41 Withdraw
  C42 Update
  C43 Validate
  C44 Annotate
  C45 Manage Inform Object

C64 Manage Actor
  C71 Manage Function
  C72 Manage Policy
  C73 Manage Quality Parameter

C74 Collaborate
  C75 Exchange Information
  C76 Converse
  C77 Find Collaborator
  C78 Author Collaboratively

C79 Manage DL
  C80 Manage Content
  C85 Manage User
  C90 Manage Functionality
  C92 Manage Quality
  C93 Manage Policy Domain

C94 Manage & Configure DLS
  C95 Manage DLS
  C104 Configure DLS
The Policy domain
Categorization of Policies

- Policy
  - Policy by characteristic
  - Policy by scope
    - System Policy
      - Support Policy
      - Connectivity Policy
      - Resource Management Policy
      - Change Management Policy
    - Functionality Policy
      - Security Policy
      - Access Policy
      - User Policy
    - Digital Rights
      - Acceptable User Behaviour Policy
    - Acceptable User Behaviour Policy

- Policy by compliance
  - Enforced Policy
    - Extrinsic Policy
  - Voluntary Policy
    - Intrinsic Policy
  - Implicit Policy
    - Prescriptive Policy
  - Explicit Policy
    - Descriptive Policy

- Policy by expression
  - Policy by context
  - Policy by application
The Quality Domain
The Architecture Domain
The Digital Library Development Framework

Reference Model

Reference Architecture

Concrete Architecture

Implementation

Related Work
- Best Practices
- Research
- Protocols
- Standards
- Specifications
- Off the shelf Components and Systems

Input
- Goals
- Requirements
- Motivations
- Market

abstract

concrete

accounts for

basis for

constrains

guides

considers

considers

use
Thank you

www.dlorg.eu