DL.org

Digital Library Reference Model In a Nutshell

www.dlorg.eu

COOPERATION

Digital Library Reference Model - In a Nutshell

This booklet is abstracted and abridged from "The Digital Library Reference Model", D3.2b DL.org Project Deliverable, April 2011

Authors

L. Candela[1], G. Athanasopoulos[2], D. Castelli[1], K. El Raheb[2], P. Innocenti[3], Y. Ioannidis[2], A. Katifori[2], A. Nika[2], G. Vullo[3], S. Ross[4]

[1] National Research Council of Italy
[2] University of Athens
[3] University of Glasgow
[4] University of Toronto

Contributing Authors

The following authors have contributed to the creation of the DELOS Digital Library Reference Model, a predecessor of "The Digital Library Reference Model"
L. Candela; D. Castelli; N. Ferro; Y. Ioannidis; G. Koutrika; C. Meghini; P. Pagano; S. Ross; D. Soergel; M. Agosti; M. Dobreva; V. Katifori; H. Schuldt

Acknowledgements

This work has been partially supported by DL.org (December 2008-February 2011), a Coordination and support action, received funding from the Commission of the European Union (EC) under the 7th Framework Programme ICT Thematic Area "Digital libraries and technology-enhanced learning" through the EC's Cultural Heritage and Technology Enhanced Learning Unit.

Designed and editorial work by Trust-IT Services Printed by Osmotica

Contents

1. Introduction

The Digital Library universe is a complex framework bringing together many disciplines and fields, spanning data management, information retrieval, library sciences, document management, information systems, web image processing, artificial intelligence, human-computer interaction and digital curation. The Digital Library universe is also an interplay of professional roles, encompassing cataloguing and curating, defining, customising and maintaining the Digital Library and its services, as well as developing and customising software. Such complexity and diversity in terms of approaches, solutions and systems has driven the need for common foundations that foster best practices and help focus further advancement in the field. The Digital Library Reference Model aims at contributing to the creation of such foundations. It is the result of a collective understanding on Digital Libraries that has been acquired by European research groups under the umbrella of the European funded DELOS Network of Excellence on Digital Libraries, as well as the international scientific community active in the field of Digital Libraries. The outcomes of DELOS have been taken forward by DL.org, a project funded by the Cultural Heritage and Technology Advanced Learning Unit of the Information Society Directorate-General of the European Commission, working in synergy with a team of international experts in the field to enhance and extend the Reference Model.

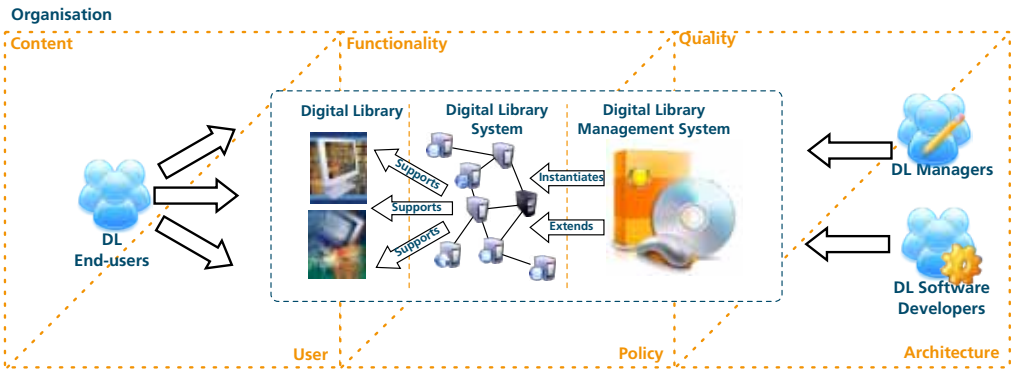
The Digital Library Reference Model is a conceptual framework aimed at capturing significant entities and their relationships in the digital library universe with the goal of developing a concrete model of it. This conceptual framework can be exploited for coordinating approaches, solutions and systems development in the digital library area. In particular, it is envisaged that in the future Digital Library 'systems' will be described, classified and measured according to the key elements introduced by this model.

Behind the Reference Model's efforts there has been the driving force of The Digital Library Manifesto, laying down the main notions characterising the Digital Library universe in rather abstract terms.

1.1 The Digital Library Manifesto – In Brief

The Digital Libraries field is highly multidisciplinary and this has created several conceptions of what a Digital Library is, each one influenced by the perspective of the primary discipline of the conceiver. Although the field cannot be captured by a simple definition, a comprehensive representation encapsulating all potential perspectives was judged as required. This led to the drafting of *The Digital Library Manifesto*, whose aim is to set the foundations and identify the cornerstone concepts within the universe of Digital Libraries, facilitating the integration of research and proposing better ways of developing appropriate systems.

The Manifesto first presents an examination of the three types of relevant 'systems' in this area: Digital Library, Digital Library System, and Digital Library Management System, describing them as follows. **Digital Library** (DL) is a potentially virtual organisation, that comprehensively collects, manages and preserves for the long depth of time rich digital content, and offers to its target user communities specialised functionality on that content, of defined quality and according to comprehensive codified policies. It is the final 'system' actually perceived by the end-users as being the digital library. **Digital Library System** (DLS) is the deployed and running software system that implements the DL facilities. **Digital Library Management System** (DLMS) is the generic software system that supports the production and administration of DLSs and the integration of additional software offer-ing more refined, special-ised or advanced facilities.



It then individuates the main concepts characterising the above systems, classifying them in a number of *domains*, each of them representing a particular aspect of the Digital Library universe. These domains are: (i) **Organisation** – represents the social arrangement characterising the expected DL service. It is a super domain that comprises the remaining six domains that actually characterise the service; (ii) **Content** – represents the information managed; (iii) **User** – represents the actors interacting with the system; (iv) **Functionality** – represents the facilities supported; (v) **Policy** – represents the rules and conditions, including digital rights, governing the operation of the whole; (vi) **Quality** – represents the aspects of digital library systems to be considered from a quality point of view; (vii) **Architecture** – represents the software (and hardware) constituents concretely realising the whole.

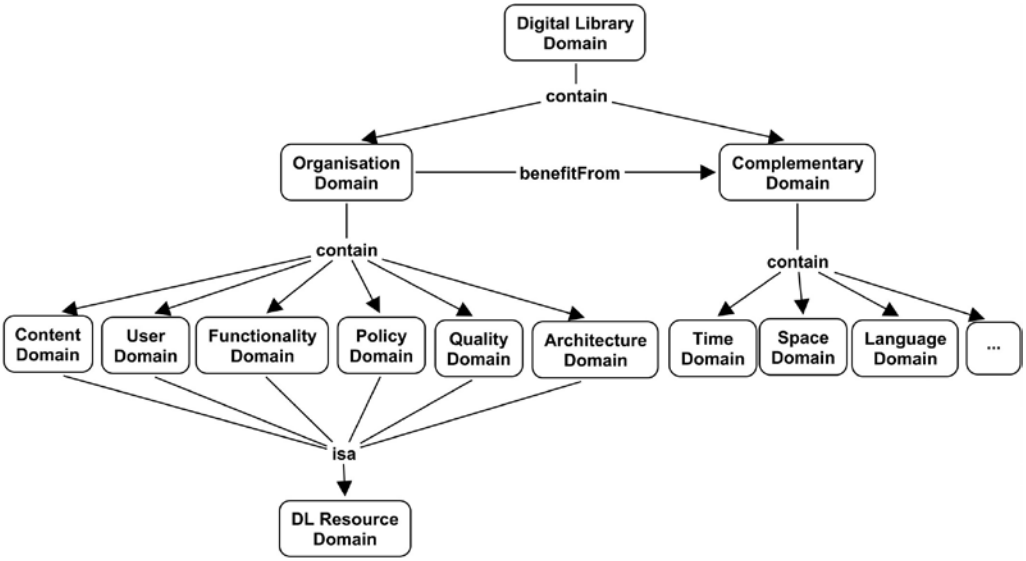
It also introduces the main roles that actors may play within digital libraries, namely: (i) **Digital Library End-Users**: the ultimate clients that the Digital Library is designed to serve. End-users are divided into *Content Creators* – 'producers' of the Digital Library Content; *Content Consumers* – 'clients' of the Digital Library Content; and *Digital Librarians* – the 'curators" of Digital Library Content; (ii) **Digital Library Managers**: 'drivers' of the Digital Library service, that is, the actors needed to put the planned service in place. They are further divided into *DL Designers* – the actors requested to characterise the Digital Library service before it is deployed; and *Digital*

Library System Administrators – the actors assigned to deploy the Digital Library System needed to implement the Digital Library Designers plan; (iii) **Digital Library Software Developers**: implementers of the software parts needed to create the Digital Library service.

Finally, the Manifesto presents a plan for laying down a comprehensive characterisation of the digital library universe having the just described systems, domains and roles of actors as solid foundations. The envisaged characterisation is based on different artefacts capturing the universe at diverse levels of abstraction from the very abstract one, i.e. the Digital Library Reference Model, to the very concrete one, i.e. the implementation. The remainder of this booklet focuses on the Digital Library Reference Model, i.e. a model consisting of a set of unifying concepts, axioms and relationships characterizing the digital library domain independently of specific standards, technologies, implementations or other concrete details. It briefly presents the constituent domains concepts are organized in and describes their rationale. Overall, this model consists of 200+ concepts and 50+ relations.

2. Constituent Domains

The Reference Model organizes the multi-faced aspects of the digital library universe into a hierarchy of *domains*, i.e., named groups of concepts and relations, each modelling a certain aspect of the systems of the universe. Domains may rely on each other and constitute orthogonal areas intended to capture the different aspects of the whole.



The ***Digital Library Domain***, which comprises all the elements needed to represent the three systems of the digital library universe, is divided into two main classes: ***Organisation Domain*** and ***Complementary Domain***. The *Organisation Domain* stems from the Organisation core concept and it is conceived to represent the main settings for characterising the DL service, the aspects that are specific to the digital library universe. It contains the following sub-domains, in full correspondence with the remaining core concepts identified in the Digital Library Manifesto: ***Content Domain, User Domain, Functionality Domain, Policy Domain, Quality Domain, Architecture Domain.*** Each of such domains focuses on a particular aspect characterising the digital library universe. However, independently of the specific aspect each domain is dedicated to, there

are some commonalities that these aspects share and these have been captured by the ***DL Resource Domain***, described below.

The *Complementary Domain* contains all the other domains, which, although they do not constitute the focus of the digital libraries and can be inherited from existing models, are nevertheless needed to represent the DL service.

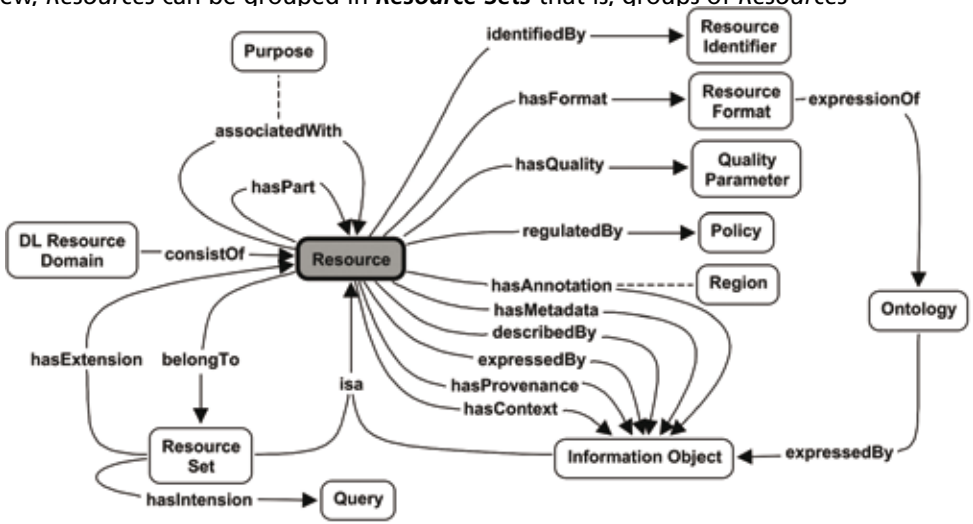
2.1 Resource Domain

This domain captures the commonalities shared by all entities and relationships that are managed in every "digital library". The most general concept of the *DL Resource Domain* is ***Resource***, which captures the characteristics of any Digital Library entity. Instances of the concept of *Resource* in the Digital Library universe are *Information Objects* in all their forms, *Actors*, *Functions*, *Policies*, *Quality Parameters* and *Architectural Components*. These instantiate the main concepts in their respective domain, thus every *Domain* consists of *Resources*, and *Resources* are the building blocks of all the *Digital Library Domains.*

All the different types of *Resources* share many characteristics and ways in which they can be related to other Resources.

Each *Resource* is: *(i)* Identified by a ***Resource Identifier***; *(ii)* arranged or laid out according to a ***Resource Format*** – such a format may be drawn from an Ontology to guarantee a uniform interpretation; it can be arbitrarily complex and structured, because *Resources* may be com-posed of smaller *Resources* and linked to other *Resources* (<*associatedWith*>); *(iii)* characterised by various ***Quality Parameters,*** each capturing how the resource performs with respect to some attribute; (iv) regulated by ***Policies*** governing every aspect of its lifetime; (v) expressed by an ***Information Object*** (such as a *Policy* set down in a text or a flowchart); (vi) described by or commented on by an *Information Object*, especially by those dedicated to record *Metadata, Annotations, Context,* or *Provenance.*

From an organisational point of view, *Resources* can be grouped in ***Resource Sets*** that is, groups of *Resources* to be considered as a single entity for certain management or application purposes. Examples of a *Resource Set* in the various domains are *Collection* in the *Content Domain* or *Group* in the *User Domain*. Every Resource Set is characterised by an intension (<*hasIntension*>) and an extension (<*hasExtension*>). The former is a criterion underlying the grouping and corresponds to a *Query*, that is, every *Query* identifies a set



of *Resources*. The way this criterion is expressed can range from the explicit enumeration of all the objects intended to be part of the group to logical expressions capturing the characteristics of the *Resources* intended to be part of the group. The latter is the concrete set of *Resources* matching the intension, that is, the set of Resources belonging to (<*belongTo*>) the *Resource Set*. These characteristics are implemented differently in diverse systems, leading to scenarios that range from static to highly dynamic ones.

Modelling the characteristics shared by all the main entities of the digital library universe at a high level of abstraction and representing more specific entity types by inheriting the shared characteristics lead to a sophisticated and concise model, to efficient implementations, and uniform user interfaces. The advantages of this modelling approach can be transformed into innovative system features and implementations. For example, unified mechanisms for handling relations and functions that apply to all resource types and unified search facilities for seamless discovery of the various entities available in a Digital Library can be envisaged.

2.2 Content Domain

This domain represents all the entities managed by the Digital Library 'systems' to meet the information needs of their users.

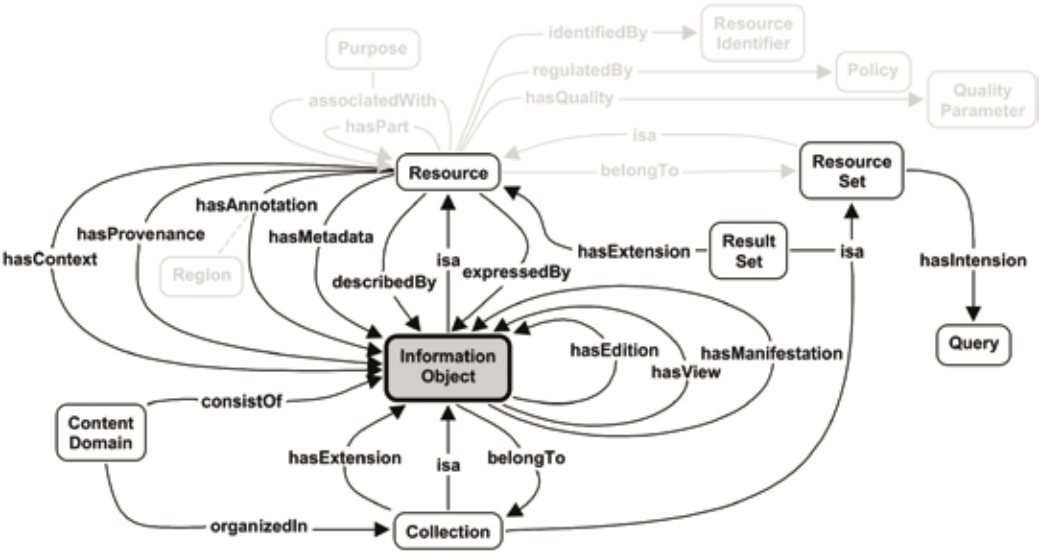
The most general concept in the Content Domain is ***Information Object***. An *Information Object* represents any unit of information such as text documents, images, sound documents, multimedia documents and 3D objects, including games and virtual reality documents, as well as data sets and databases. *Information Object* also includes composite objects and *Collections* of *Information Objects*. As an *Information Object* is a *Resource*, it inherits all its features.

Information Objects can be grouped into ***Collections*** (<belongTo>), that is, special type of *Resources* which are themselves *Information Objects* and inherit all *Information Objects'* features, for example, they can be annotated. Collections are a specialisation of the *Resource Set* concept. They are characterised by an intension *(<hasIntension>)* – the *Query* capturing the criterion underlying the group – and an extension *(<hasExtension>)* – the set of *Information Objects* matching the intension. Another specialisation of the *Resource Set* concept usually associated with the *Content Domain* is the **Result Set**. In traditional digital libraries this is the set of documents that are retrieved by issuing a **Query**. In this context it represents the set of *Resources*, with no constraints on their type, resulting from a *Query.*

Information Objects can acquire specialisations depending on various aspects. All of them are expected to be captured by relying on relations between Information Objects. One of these aspects is the level of abstraction at which they are specified. This leads to an abstract *Information object* by level of abstraction concept, which is a container or placeholder to be specialised using any of several models. For example, it can be useful to represent the IFLA FRBR model.

Information objects can also be specialised by the predominant role they play in their relationship to other objects; the class *Information object by relationship* is the abstract conceptual container for the classes these objects give rise to, namely: (i) *Primary Information Object*, an *Information Object* that stands on its own, such as a book or a data set; (ii) **Metadata** object, an *Information Object* whose predominant purpose is to give information about a 'target' *Resource* (usually, but not always, a *Primary Information Object*); (iii) **Annotation** object, an *Information*

Object whose predominant purpose is to annotate a 'target' *Resource* (or a **Region** of it). Examples of such *Annotation Objects* include notes, structured comments, and links. *Annotation Objects* assist in the interpretation of the target *Resource*, or give support or objections or more detailed explanations.



A distinguishing characteristic of this model with respect to most Digital Library models or de facto standards is that an *information object* is not born as (say) *Metadata* or as *Annotation*, but becomes such by virtue of playing a certain role in relation to other information objects. A typical case arises for a piece of text; it is primarily a piece of text, and becomes an annotation only when it is linked to a certain Resource in a certain way. In other words, the long-standing issue of whether annotations are content or metadata is an ill-posed question.

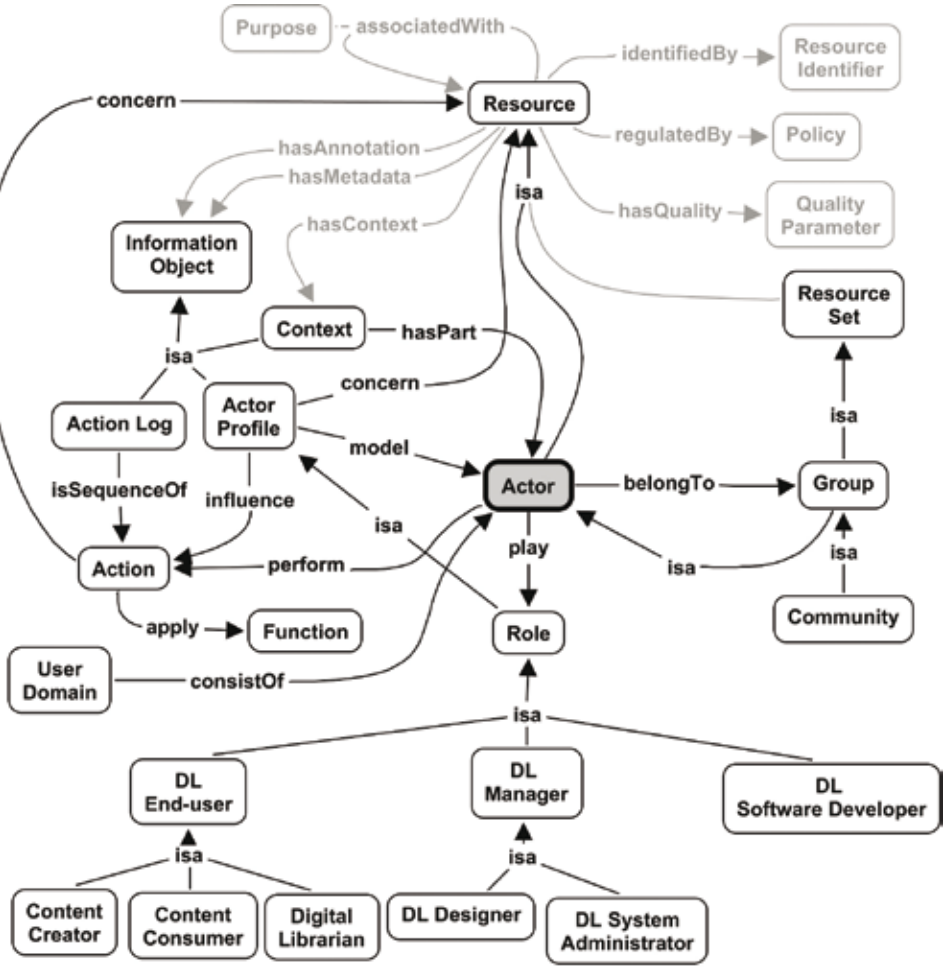
Finally, various types of Information Objects can be distinguished although all of them are instances of the same concept. Possible dimensions are: (i) by the type of representation or encoding: e.g. Information Objects encoded in some natural form directly interpretable by human, text in natural language, images, sounds, etc; Information Objects encoded in a formal structure, such as database tables, formal entity-relationship statements, ontologies in formal terms; (ii) by the relationship to real world objects: e.g. born digital: information objects such as born digital texts or digital camera images, which are the real world objects themselves and do not correspond to any other real world objects; Information objects produced by digitisation of non-digital information objects, such as digitised versions of ancient manuscripts; Information object representing *Metadata*, such as the descriptive information of the Mona Lisa, describing a real-world object, whether the latter is digital or not, or represented in the Digital Library or not.

2.3 User Domain

This domain represents all the entities that interact with any Digital Library 'system', that is, humans and inanimate entities such as software programmes or physical instruments. Exemplars of inanimate entities include a subscription service offered by a university to its students, which provides access to the contents of an external or another Digital Library. Inclusion of hardware and software into the potential users of digital libraries marks a shift away from other Digital Library models and reflects a broader concept of 'digital library'.

To capture these extended semantics, we use the concept of ***Actor*** as the dominant concept in this domain. Being a *Resource*, the *Actor* concept inherits all key charac-teristics of the former. An ***Actor Profile*** is used to model an *Actor*. Every *Actor* interacts with the Digital Library, Digital Library System or Digital Library Management System by performing certain ***Action(s)***.

The *Actor Profile* is an *Information Object* that concerns *Resources* and essentially models an *Actor* by capturing a large variety of the *Actor's* potential characteristics. It enables the *Actor* to interact with the 'system' as well as with other *Actors* in a personalised, customised way. Not only does it serve as a representation of *Actor* in the system but also essentially captures the *Policies* and *Roles* that govern which *Functions* are allowed on which *Resources* by the *Actor*. For example, a particular instance of *Actor* may be entitled to *Search* within particular *Collections* and to *Collaborate* with particular other *Actors*. The characteristics captured in an *Actor Profile* vary depending on the type of *Actor*, whether human or non-human, and may include: demographic information, such as age, residence or location for humans and operating system, web server edition for software components, educational information such as highest degree achieved, field of study for humans, and preferences, such as topics of interest, pertinent for both human and software *Actors* that interact with the Digital Library.



An *Actor* may play a different ***Role*** at different times, a conception which marks a significant shift away from traditional approaches, where there are typically strong dependencies between *Roles* and *Actors* and an *Actor* can typically play one *Role*. Among *Actor Roles*, important categories are ***End-user***, ***Digital Library Manager***, and ***Digital Library Software Developer***. Each of these *roles* plays a complementary activity along the 'system' life-cycle. *End-user* leverages digital library facilities for providing, consuming and managing digital library content. It is further subdivided

into the concepts of ***Content Creator***, ***Content Consumer*** and ***Digital Librarian***, each of which usually has a different perspective on the Digital Library. For instance, a *Content Creator* may be a person that creates and inserts his or her own documents in the Digital Library or an external programme that automatically converts documents to digital form and uploads them to the Digital Library. *Actors* in the role of *Digital Library Manager* leverage Digital Library Management System facilities to define, customise and maintain the digital library service. It is further subdivided into ***Digital Library Designers***, who define, customise and maintain the service – and ***Digital Library System Administrators***, who leverage Digital Library Management System facilities to deliver and operate the Digital Library Service foreseen. Finally, *Digital Library Software Developers* leverage Digital Library Management System facilities to create and customise the constituents of the Digital Library System and Digital Library Management System. Inclusion of this broad understanding of actor roles into the potential users of Digital Libraries marks a major shift away from other Digital Library models that focus on the *End-user* part only.

Finally, an *Actor* may be part of a ***Group***. A *Group* represents a set of *Actors* which exhibits cohesiveness to a large extent and can be considered as an *Actor* with its own profile and identifier. Members of a *Group* inherit (some of) the characteristics from the *Group*, such as interests and *policies*, but they may have additional characteristics as described in their individual *Actor's* profile. A particular sub-class of *Group* is ***Community***, which refers to a social group of humans with shared interests. In human *communities*, intent, belief, resources, preferences, needs, risks and several other conditions may be present and common, affecting the identity of the participants and the extent of cohesiveness.

2.4 Functionality Domain

This domain is one of the richest and most open-ended dimensions of the world of digital libraries, as it captures all the processing that can occur on *Resources* and actions that can be observed by *Actors* in a Digital Library, Digital Library System or Digital Library Management System.

The most general functionality concept is ***Function***, that is, a particular processing task that can be realised on a *Resource* or *Resource Set* as the result of an activity of a particular *Actor*. It is worth noting that this description of a *Function* is based on the generalised concepts of *Actor*, capturing not only human users but also inanimate entities, and of *Resource*, representing all entities involved in or influenced by a Digital Library, Digital Library System or Digital Library Management System. Hence, this description lends a new perspective to the *Functions* of this domain. For instance, not only can a human *Actor Search* the contents in a digital library (*Information Objects*), but also for other *Actors*; a programme can *Search* for offered *Functions*, and so forth. Each *Function* is itself a *Resource* in this model and thus inherits all the characteristics of the former.

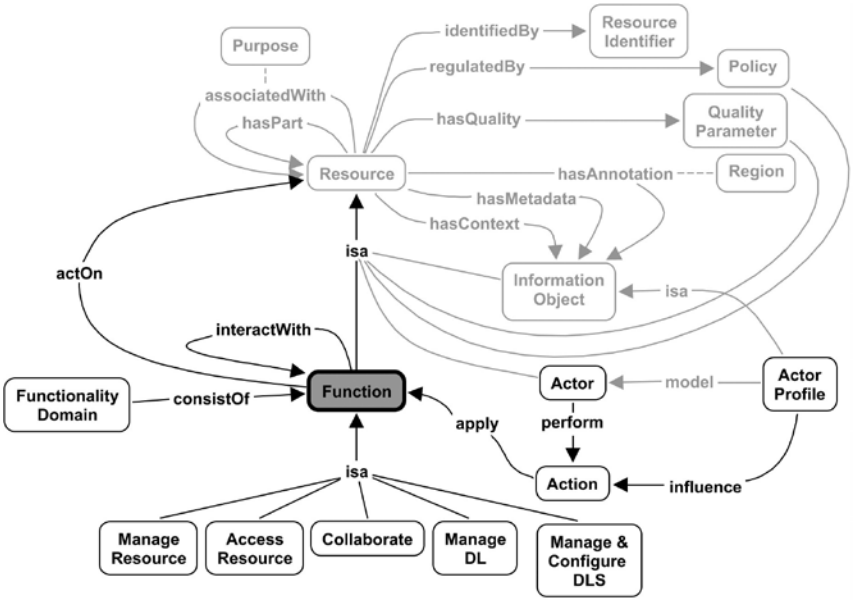
The broad scope of the *Function* concept precludes enumerating and predicting all the different types and 'flavours' of *Functions* that may be included in a Digital Library, Digital Library System or Digital Library Management System. Each one may have its own set of *Functions* depending on its objectives or its intended *Actors*. Therefore, the *Function* concept is specialised into five sub-concepts that still represent quite general classes of activities. The first three types of *Functions* (*Manage Resource*, *Access Resource*, *Collaborate*) accommodate activities related to the prime actions, which are performed by the digital library *Actors* – namely *End-user*.

Manage Resource includes all activities related to creating new *Resources* and making them available through the Digital Library, deleting old *Resources* from it, and updating existing ones. General management *Functions* that are applicable on all *Resources* include the creation, submission, withdrawal, update, preservation, validation and annotation. In addition to these general functions, other Functions result when dealing with specific kind of *Resources*, e.g. *Information Objects*, *Actors*, *Policy*. Given their basic role, two of the *Manage Resource Functions* merit detail: *Manage Information Object* and *Manage Actor*. ***Manage Information Object*** is the family of *Manage Resource Functions* conceived to capture those dedicated to *Information Objects*. This family contains *Functions* supporting authoring and dissemination as well as a rich array of actions dedicated to Information Object processing. ***Manage Actor*** is the family of *Manage Resource Functions* designed to capture *Functions* necessary for the management of individual *Actors*, including their registration or subscription, their log-in and profiling.

The second type of prime action expected to be performed by *End-user* deals with accessing the digital library offering. ***Access Resource*** encompasses all activities related to requesting, locating, retrieving, browsing, and representing *Resources*. The key characteristic of the *Access Resource* concept is that it represents *Functions* that do not modify the Digital Library (Digital Library System and Digital Library Management System as well) but identify *Resources* to be sensed by *Actors* or possibly further exploited by other *Functions*. Hence, the central *Access Resource* function is *Discover*, which acts on *Resource Sets* to retrieve desired *Resources*. The third type of prime action expected to be performed by *End-user* deals with designing the digital library service as a collaborative working environment. ***Collaborate*** is the family of *Functions* capturing all activities that enable multiple *Actors* to work together on top of a Digital Library to achieve a common goal. It explicitly captures

the main *Functions* that fall into this domain including basic facilities, such as collaborative authoring via *Author Collaboratively*, and facilities promoting the collaboration, e.g. co-workers discovery via *Find Collaborator*.

The remaining two specialisations of the *Function* concept encompass all activities related to the 'system' as a whole and its management. These specialisations are *Manage DL* and *Manage & Configure DLS*. They are oriented to support the activities of the *Actors* needed to operate the digital library service – mainly, *Digital Librarians* and *DL Managers* as well as *DL Software Developers* – and are expected to be supported by: the Digital Library – for day-to-day management (*Manage DL*) and *Digital Library Management System* – for long-term management (*Manage & Configure DLS*). ***Manage DL*** includes a wide variety of *Functions* that support the day-to-day management of the overall DL service. Because of this, it includes facilities for revising every aspect of the service from *Content* (e.g., *Collection management*) and User (e.g., *Group management*) -related characteristics to *Functionality*, *Policy* and *Quality* ones. These *Functions* are mainly associated with the role of *Digital Librarian*. However, part of them can be associated with the role of *DL Designer*. ***Manage & Configure DLS*** contains *Functions* serving the *DL Manager,* in particular, the *DL System Administrator* in terms of setting up, configuring and monitoring the digital library service from a physical point of view, that is, deploying the Digital Library System needed to implement and support the Digital Library as foreseen.



Functions realise what is usually called a 'business process' which is in the service of meeting specific 'business requirements' that meet a 'stakeholder need'. As the *Functionality Domain* is among the most dynamic of all fundamental *Domains* in the Digital Library Universe, the DL.org Reference Model represents only a sub-set of *Functions, with* special emphasis on the most critical ones, that is, *Functions* available in most of the existing Digital Library Systems and needed to support interaction with the intended clients or *Functions* expected by Digital Library Management Systems to deploy and operate the service.

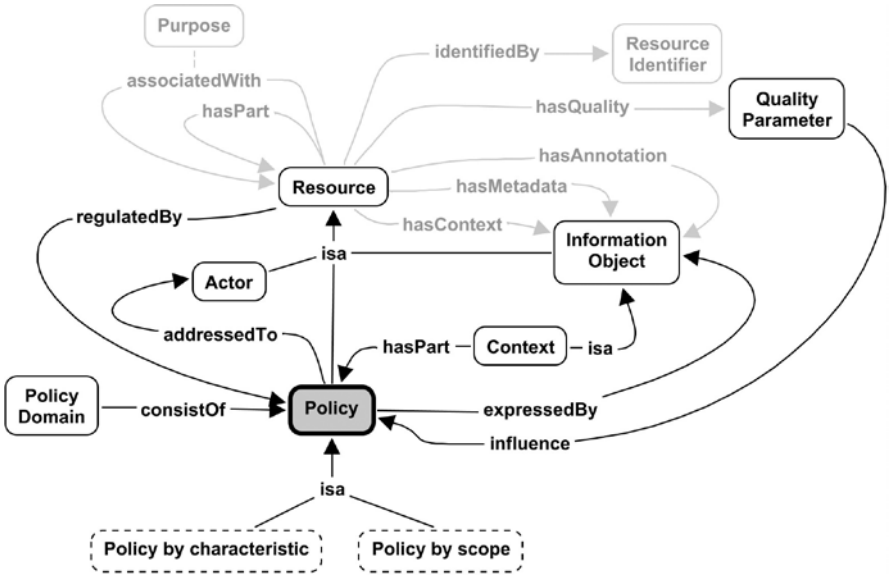
2.5 Policy Domain

This domain represents the set of conditions, rules, terms or regulations governing the operation of any digital library 'system', that is, Digital Library, Digital Library System and Digital Library Management System. Policy at large governs the operation of any kind of 'system' including our society or the Institution or Organisation that sets up the Digital Library. Policies are always addressed to defined *Actors*. This domain is, by definition, very broad and dynamic. The representation provided by this model does not purport to be exhaustive, especially with respect to the myriad of specific rules each Institution would like to model and apply. The *Policy domain* captures the minimal set of relationships connecting it to the rest and presents the kind of rules that are considered as most critical in the Digital Library universe.

The most general policy concept is ***Policy***, the entity regulating the existence of a *Resource* with respect to a certain management point of view. Each *Policy* is itself a *Resource* in this model and thus inherits all the characteristics of the former.

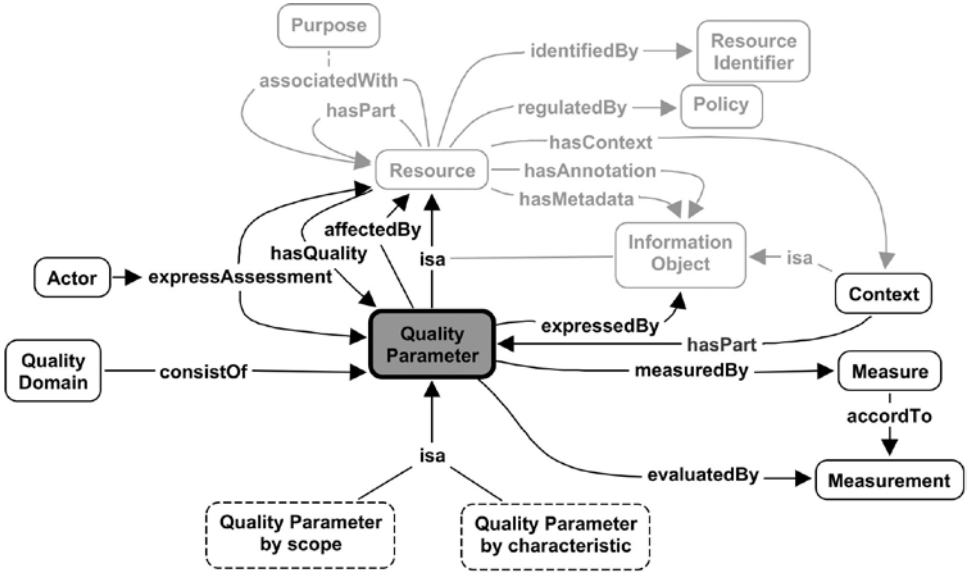
Policy is actually a class of various types of policies. For the purpose of this model, two abstract and orthogonal conceptual containers have been identified, that is, ***Policy by characteristic*** and ***Policy by scope***. Policy by characteristic is further specialised into eight sub-classes, each presenting a bipolar quality a *Policy* might have: ***Extrinsic Policy*** vs. ***Intrinsic Policy***; ***Implicit Policy*** vs. ***Explicit Policy***; ***Prescriptive Policy*** vs. ***Descriptive Policy***; ***Enforced Policy*** vs. ***Voluntary Policy***. Understanding the characteristics of a specific *Policy* helps to express it better and to clarify requirements at all levels across the boundaries of the three 'systems': Digital Library, Digital Library System and Digital Library Management System.

Policy by scope is further specialised into various classes, each representing a particular *Policy* with respect to (a) the system as a whole, for example, *Resource Management Policy*; (b) a certain domain, for example., *User Policy* or *Content Policy*. In some cases a *Policy* actually serves the needs of multiple domains, for example, *Access Policy* is both a *User Policy* and a *Functionality Policy.* It is important to remember that the model is extensible and does not intend to form an exhaustive list but rather a sample capturing some of the most important *Policies* governing the Digital Library universe. Among them, a special role is occupied by the *Digital Rights Management Policy* and *Digital Rights*.



2.6 Quality Domain

This domain represents aspects that permit considering any digital library 'system' from a quality point of view, with the goal of judging and evaluating them with respect to specific facets. Any digital library 'system' tenders a certain level of *Quality* to its *Actors* that can be either implicitly agreed, that is, *Actors* simply have an understanding of what *Quality Parameters* are guaranteed, or explicitly formulated, in that there is a Quality of Service (QoS) agreement.



The most general quality concept is ***Quality Parameter***, that is, the entity expressing the different facets of the *Quality Domain* and providing information about how and how well a *Resource* performs with respect to some viewpoint. *Quality Parameters* express an assessment by an *Actor*, whether human or not, of the *Resource* under considera-tion. The *Quality Parameters* can be evaluated according to different *Measurements*, which provide alter-native procedures for assessing dif-ferent aspects of each *Quality Parameter* and assigning it a value. *Quality Parameters* are actually expressed by a *Measure*, which represents the value assigned to *a Quality Parameter* with respect to a selected *Measurement*.

In this model each *Quality Parameter* is itself a *Resource*, thus inheriting all its characteristics.

The *Quality Domain* is, by definition, very broad and dynamic, extensible with respect to the myriad of specific quality facets each Institution would like to model. These parameters are grouped according to the *Resource* under examination, that is, ***Quality Parameter by scope***, and to the characteristics of the *Measurement*, that is, ***Quality Parameter by characteristic***. *Quality Parameter by scope* is further specialised in: ***Generic Quality Parameter*** – it applies to any kind or most kinds of *Resources*; ***Content Quality Parameter*** – it applies to *Resources* in the *Content Domain*, namely *Information Objects*; ***Functionality Quality Parameter*** – it applies to *Resources* in the *Functionality Domain*, namely *Functions*; ***User Quality Parameter*** – it applies to *Resources* in the *User Domain*, namely *Actors*; ***Policy Quality Parameter*** – it applies to *Resources* in the *Policy Domain*, namely *Policies*; ***Architecture Quality Parameters*** – it applies to *Resources* belonging to the *Architecture Domain*, namely *Architectural Components*. It is important to note that this grouping is made from the perspective of the *Resource* under examination, that is, the main object under assessment. In any case, the *Actor*, understood as the active subject who expresses the assessment, is always taken into consideration and explicitly modelled, since he/she is an integral part of the definition of *Quality Parameter*. Therefore, the ***User Satisfaction*** parameter has been grouped under the *Functionality Quality Parameter* because it expresses how much an *Actor* (the subject who makes the assessment) is satisfied when he/she/it uses a given *Function* (the object of the assessment).

2.7 Architecture Domain

This domain includes concepts and relationships characterising the two software systems playing an active role in the Digital Library universe, that is, Digital Library Systems and Digital Library Management Systems. The importance of this fundamental con-cept has been large-ly underestimated in the past. Having a clear architectural understanding of the software sys-tems implementing the Digital Library u-niverse offers guide-lines on pragmatic set-up of a Digital Library as a whole. In particular, it of-fers insights into: (a) how to develop new systems, by maxi-mising sharing and reuse of valuable assets to minimise the development cost and time-to-market; and (b) how to improve current systems by promoting the adoption of suitable, recognisable, and widely accepted patterns to simplify interoperability issues.

The most general concept in the *Architecture Domain* is ***Architectural Component***, that is, a significant system component. Thus, for the purposes of this Reference Model, the architecture of a software system (at a given point) is defined as the organisation or structure of its *Architectural Components* interacting with each other through their interfaces (**Interface**). These components may in turn be composed of smaller and smaller components; however, different *Architectural Components* may be incompatible with each other, that is, cannot co-exist in the context of the same system. When using the term 'component' the software industry and the literature refer to many different concepts. Here, we use the term 'component' to denote an encapsulated part of a system, ideally a 'non-trivial', 'nearly independent', and 'replaceable' part of a system that fulfils a clear function in the context of a well-defined architecture.

Each *Architectural Component* is a *Resource*, thus it inherits the *Resource's* characterising aspects, for example, it is uniquely identified. Like any *Resource*, components have *Metadata* (***Component Profile***) which are expected to capture fundamental information for managing these kinds of *Resource* including the implemented or supported *Functions*, the implemented *Interfaces*, their governing *Policies*, and the *Quality Parameters* characterising them.

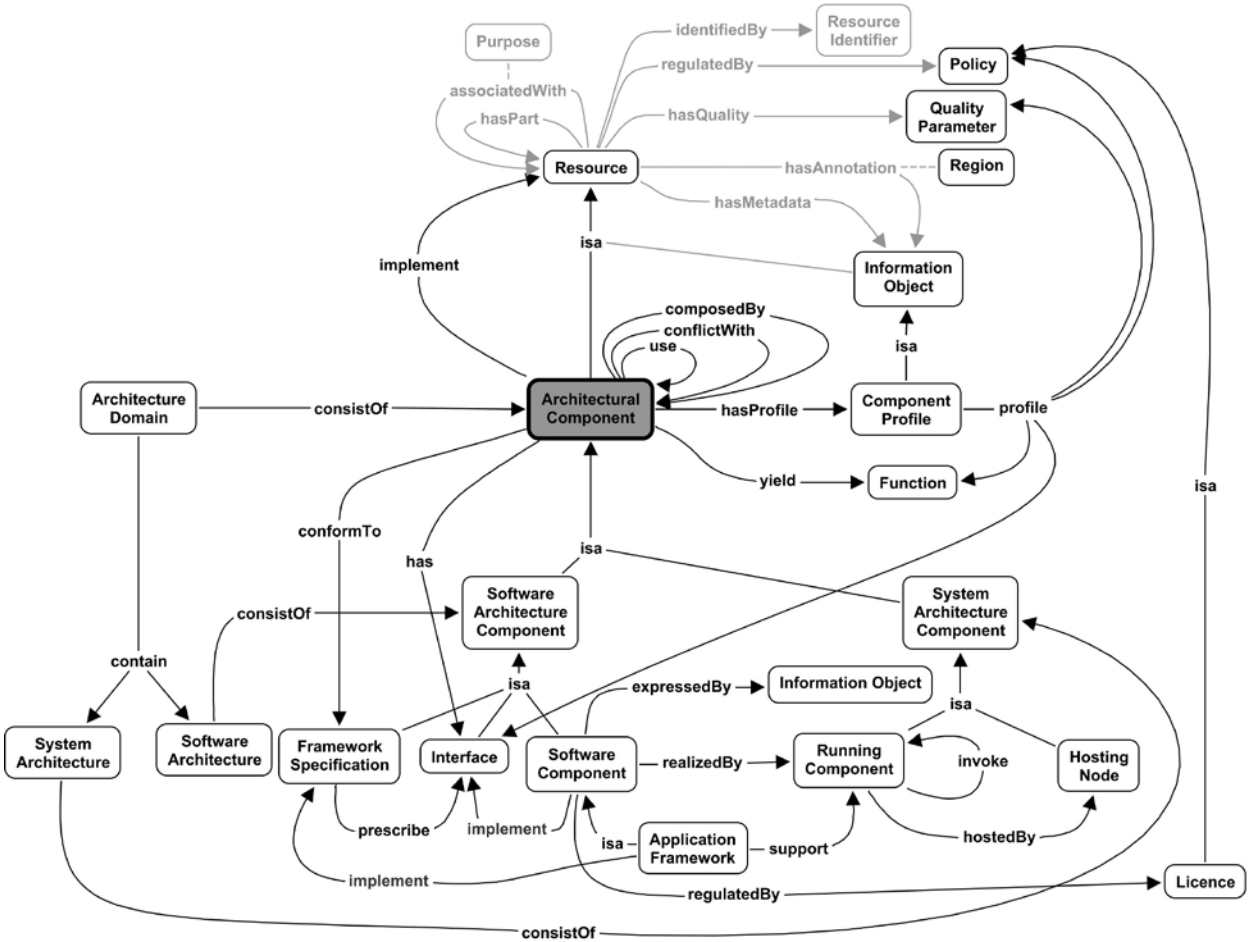
Architectural Components interact through a ***Framework Specification*** and are conformant to it. This framework prescribes the set of *Interfaces* to be implemented by the components and the protocols governing how components interact with each other. *Architectural Components* are classified into ***Software Architecture Components*** and ***System Architecture Components***. These classes are used to describe the ***Software Architecture*** and the ***System Architecture*** of a software system respectively, where the former captures the organisation of the programs a software system consists of, while the latter captures the organisation of the processes and running units an operating software system consists of.

Software Architecture Components are realised by ***Software Components***. A *Software Component*, encapsulates the implementation of a portion of a software system and is regulated by particular *Policies* (***Licenses***). Moreover, it is represented by an *Information Obje*ct. Thus, the *Resource* representing the *Software Component* inherits the *Information Object's* characterising aspects, for example, it can be enriched through *Metadata* and *Annotations*. Exemplars of *Software Architecture Components* are software packages implementing a specific *Function*, software artefacts supporting the implementation of a specific *Function*, for example a Relational Database Management System (RDBMS).

System Architecture Components are realised by ***Hosting Nodes*** and ***Running Components***. A *Hosting Node* represents the (virtual) hardware environment hosting and running *Software Components*. A *Running*

Component represents a running instance of a *Software Component* active on a *Hosting Node*. Exemplars of *System Architecture Components* are servers that can host one or more of the Digital Library System processes or running units, an operational Web Service partaking to the System Architecture of a Digital Library System, a deployed Relational Database Management System (RDBMS).

Overall, this modelling subsumes a 'component-based approach', that is, a kind of application development in which: (i) the system is assembled from discrete executable components, which are developed and deployed somewhat independently of one another, and potentially by different players; (ii) the system may be upgraded with smaller increments, that is, by upgrading some of the constituent components only. In particular, this aspect is one of the key points for achieving interoperability, as upgrading the appropriate constituents of a system enables it to interact with other systems; (iii) components may be shared by systems; this creates opportunities for reuse, which contributes significantly to lowering the development and maintenance costs and the time to market; (iv) though not strictly related to their being component-based, component-based systems tend to be distributed.



3. Bibliography

Bertino, E., Casarosa, V., Crane, G., Croft, B., Del Bimbo, A., Fellner, D., Fiander, P., Fox, E., Fuhr, N., Ioannidis, Y., Koch, T., Larsen, R., Manson, P., Marchetti, P. G., Mylopoulos, J., Neuhold, E., Paolini, P., Papazoglou, M., Romary, L., Schek, H.-J., Smeaton, A., Thanos, C., Wactlar, H. (2001). *Digital Libraries: Future Directions for a European Research Research Programme*. San Cassiano. DELOS.

Borgman, C. (1999). What are digital libraries? Competing visions. *Information Processing and Management*, 35 (3), 227-243.

Borgman, C. (2010). Scholarship in the Digital Age: Information, Infrastructure, and the Internet. The MIT Press.

Bush, V. (1945). As We May Think. Atlantic Monthly (176), 101-108.

Candela, L., Athanasopoulos, G.; Castelli, D.; El Raheb, K.; Innocenti, P.; Ioannidis, Y.; Katifori, A.; Nika, A.; Vullo, G.; & Ross, S. (2011). The Digital Library Reference Model. DL.org Project Deliverable.

Candela, L., Castelli, D., & Pagano, P. (2011). History, Evolution and Impact of Digital Libraries. In I. Iglezakis, T.-E. Synodinou, & S. Kapidakis, E-Publishing and Digital Libraries: Legal and Organizational Issues (pp. 1-30). IGI Global.

Fox, E. A., Akscyn, R. M., Furuta, R. K., & Leggett, J. J. (1995). Digital Libraries. *Communications of the ACM*, 38 (4), 23-28.

Fox, E., & Marchionini, G. (1998). Toward a Worldwide Digital Library. *Communications of the ACM*, 41 (4), 29-32.

Griffin, S., Peters, C., & Thanos, C. (2005). Towards the new-generation digital libraries: recommendations of the NSF/EU-DELOS working groups. *International Journal of Digital Libraries*, 5 (4), 253-254.

Ioannidis, Y. (2005). Digital libraries at a crossroads. *International Journal of Digital Libraries*, 5 (4), 255-265.

Kuny, T., & Cleveland, G. (1996). The Digital Library: Myths and Challenges. Proceedings 62nd IFLA General Conference.

Lesk, M. (1999). Expanding Digital Library Research: Media, Genre, Place and Subjects. Proceedings of the International Symposium on Digital Libraries1999, ISDL'99, (pp. 51-57). Tsukuba, Ibaraki, Japan.

Licklider, J. (1965). Libraries of the Future. Cambridge: The MIT Press.

Digital Library Manifesto

DL.org (www.dlorg.eu) has mobilised professionals, educationalists and students at various stages in their academic careers mainly from Computer Science and Library and Information Science domains, to promote knowledge in digital library interoperability, best practices and modelling foundations.

DL.org Experts

Kevin Ashley (Digital Curation Centre, UK), Detlev Balzer (Independent Consultant, Germany), Tiziana Catarci (University of Rome "La Sapienza", Italy), Vassilis Christophides (University of Crete, Greece), Genevieve Clavel-Merrin (Swiss National Library, Switzerland), Antonella De Robbio (University of Padua, Italy), John Faundeen (U.S. Geological Survey Centre, U.S.), Nicola Ferro (University of Padua, Italy), Ed Fox (Virginia Tech, U.S.), Stefan Gradmann (Humboldt University, Germany), C.H.J.P. (Kees) Hendrick (Naturalis - National Museum of Natural History, The Netherlands), Sarah Higgins (Aberystwyth University, Wales, UK), René van Horik (Data Archiving and Networked Services, The Netherlands), Wolfram Horstmann (Bielefeld University Library, Germany), George Kakaletris (University of Athens, Greece), Sarantos Kapidakis (Ionian University of Corfu, Greece), Georgia Koutrika (Stanford University, U.S.), Paolo Manghi (National Research Council, Italy), Natalia Manola (University of Athens, Greece), Carlo Meghini (National Research Council, Italy), Jan Molendijk (National Library of the Netherlands, The Netherlands), Luc Moreau (University of Southampton, UK), Andreas Nürnberger (University Magdeburg, Germany), Pasquale Pagano (National Research Council, Italy), Hans Pfeiffenberger (Alfred Wegener Institute, Germany), Axel Poigné (Fraunhofer Institute Intelligent Analysis and Information Systems, Germany), Pavlos Polydoras (UBITECH, Greece), Andreas Rauber (TU-Wien, Austria), Dirk Roorda (Royal Netherlands Academy of Science, Netherlands), Robert Sanderson (Los Alamos National Laboratory, U.S.), Tefko Saracevic (Rutgers University, U.S.), MacKenzie Smith (MIT Libraries, U.S.), Thornton Staples (Smithsonian Institution, U.S.), Dagobert Soergel (University at Buffalo, U.S.), Manfred Thaller (University of Cologne, Germany), Bram van der Werf (Independent Consultant, Netherlands), René van Horik (Royal Netherlands Academy of Science, Netherlands).

DL.org Liaison Group Members

Tobias Blanke (King's College London, UK), Peter Brantley (Access for the Internet Archive, U.S.), Schubert Foo (Nanyang Technological University, Singapore), Jane Hunter (University of Queensland, Australia), Joan K. Lippincott (Coalition for Networked Information, U.S.), Clifford A. Lynch (Coalition for Networked Information, U.S.), Leonid Kalinichenko (Russian Academy of Science, Russia), Dean B. Krafft (Cornell University Library, U.S.), Carl Lagoze (Cornell University, U.S.), Ronald L. Larsen (University of Pittsburgh, U.S.), Andrea Mulrenin (Salzburg Research, Austria), Erich J. Neuhold (University of Vienna, Austria), Areti Ramachandra Durga Prasad (Indian Statistical Institute Bangalore, India), Jela Steinerová (Comenius University Bratislava, Slovakia), Shigeo Sugimoto (University of Tsukuba, Japan), Herbert Van de Sompel (Los Alamos National Laboratory, U.S.), Jens Vigen (European Organization for Nuclear Research, Switzerland), Andrew Wilson (National Archives of Australia, Australia).

DL.org Advisory Group Members

Marianne Backes (Virtual Resource Centre for Knowledge about Europe, Luxembourg), Stephen Griffin (National Science Foundation, U.S.), Geneva Henry (Rice University, U.S.).

Italian National Research Council

University of Athens

University of Glasgow | Humanities Advanced Technology & Information Institute

Trust-IT Services

DL.org 231551

ISBN: 978 889 553 4107