The Digital Library Manifesto

Leonardo Candela – Istituto di Scienza e Tecnologie dell'Informazione (ISTI-CNR)
Donatella Castelli – Istituto di Scienza e Tecnologie dell'Informazione (ISTI-CNR)
Yannis Ioannidis – University of Athens
Georgia Koutrika – University of Athens
Pasquale Pagano - Istituto di Scienza e Tecnologie dell'Informazione (ISTI-CNR)
Seamus Ross – HATII-University of Glasgow
Hans-Joerg Schek – University of Constance
Heiko Schuldt – University of Basel

DELOS: a Network of Excellence on Digital Libraries
www.delos.info
Project no.507618

DELOS

A Network of Excellence on Digital Libraries

Instrument: Network of Excellence

Thematic Priority: IST-2002-2.3.1.12
Technology-enhanced Learning and Access to Cultural Heritage

The Digital Library Manifesto
The Digital Library Manifesto

L. Candela¹, D. Castelli¹, Y. Ioannidis², G. Koutra², P. Pagano¹, S. Ross³,
H.-J. Schek⁴, H. Schuldt⁵

Abstract
The term “Digital Libraries” corresponds to a very complex notion with several diverse aspects and cannot be captured by a simple definition. A comprehensive representation of Digital Libraries encapsulating all potential perspectives is required. This has 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 ‘Digital Library Manifesto’ exploits the collective understanding that has been acquired on Digital Libraries by several previous efforts by European research groups active in the Digital Library field for many years, both within the DELOS Network of Excellence and outside, as well as by other groups around the world. Its role is one of a springboard for future work.

¹ Institute of Information Science and Technologies (ISTI)
Italian National Research Council (CNR)
Pisa, Italy

² Department of Informatics and Telecommunications
University of Athens
Athens, Greece

³ Humanities Advanced Technology and Information Institute (HATII)
University of Glasgow
Glasgow, United Kingdom

⁴ Department of Computer and Information Science
University of Konstanz
Konstanz, Germany

⁵ Database and Information Systems Group
University of Basel,
Basel, Switzerland
1 Introduction

The term ‘Digital Library’ is currently used to refer to systems that are very heterogeneous in scope and provide very different functionality. These systems span from digital object and metadata repositories, reference-linking systems, archives, and content administration systems (mainly developed by industry), to complex systems that integrate advanced digital library services (mainly developed in research environments). This “overloading” of the term ‘Digital Library’ is a consequence of the fact that there is not yet an agreement on what Digital Libraries are and what functionality is associated with them. This results in lack of interoperability and reuse of both contents and technologies. The current document attempts to put some order in the field for the benefit of its future advancement.

1.1 What is a Manifesto

According to the Merriam-Webster Dictionary, a manifesto is “a written statement declaring publicly the intentions, motives, or views of its issuer”. Similarly, according to Wikipedia, a manifesto is “a public declaration of principles and intentions”. The ‘Declaration of the Rights of Man and the Citizen’ in France in 1789 and the ‘Declaration of Independence’ in the US in 1776 are two well-known manifestos that have set the stage for the establishment of two major countries and have had major influence on the recent history of the world. The production of manifestos in subsequent centuries in fact increased; ‘The Communist Manifesto’, issued by K. Marx and F. Engels in 1848, and the ‘The Russell-Einstein Manifesto’, issued B. Russell and A. Einstein in 1955 to confront the development of weapons of mass destruction, are some of the most famous examples.

Of smaller scope and within the realm of science, there have been several manifestos as well, which have tried to draw the course for the development of particular research areas. These take more the form of declarations of axioms capturing the strategic ideas of a group of people with respect to a certain topic or field. Examples include the following:

- ‘The Third Manifesto’, from the book ‘Databases, Types, and the Relational Model: The Third Manifesto’ by H. Darwen and C. J. Date, Addison-Wesley, 2007, which proposes new foundations for future database systems\(^1\).
- ‘The Object-Oriented Database System Manifesto’, which describes the main features and characteristics that a system must have to qualify as an object-oriented database system, touching up on mandatory, optional, and even open points where the designer can make several choices\(^2\).
- ‘The Manifesto for Agile Software Development’, which attempts to uncover better ways of developing software by putting emphasis on different items than traditionally, e.g., individuals and interactions instead of processes and tools, working software instead of comprehensive documentation, and others\(^3\).
- ‘The GNU Manifesto’, by Richard Stallman, which uses as an axiom the idea that “… the golden rule requires that if I like a program I must share it with other people who like it”

\(^1\) http://www.thethirdmanifesto.com/
\(^2\) http://www.cs.cmu.edu/People/clamen/OODBMS/Manifesto/index.html
\(^3\) http://agilemanifesto.org/
produce a complete Unix-compatible software system and is given away free to everyone who can use it.

Given the current status of maturity as well as heterogeneity of the relevant research and industrial efforts, the time is right for the Digital Library (DL) field to obtain its own Manifesto. This document is *The Digital Library Manifesto*. It attempts to circumscribe the universe of Digital Libraries and set the foundations that will facilitate integration of relevant research and improvement of system development methodologies.

### 1.2 Motivation

Digital Libraries constitute a relatively young scientific field, whose life spans roughly the last fifteen years. Instrumental in the birth and growth of the field have been the funding opportunities generated by the ‘Technology Enhanced Learning; Cultural Heritage’ (formerly ‘Cultural Heritage Applications’) Unit of the Information Society Directorate-General of the European Commission and the ‘Digital Library Initiatives’ in the United States sponsored by the National Science Foundation and other agencies.

Digital Libraries represent the meeting point of a large number of disciplines and fields, i.e., data management, information retrieval, library sciences, document management, information systems, the web, image processing, artificial intelligence, human-computer interaction, and others. It was only natural that these first fifteen years were mostly spent on bridging some of the gaps between the disciplines (and the scientists serving each one), improvising on what ‘Digital-Library functionality’ is supposed to be, and integrating solutions from each separate field into systems to support such functionality, sometimes the solutions being induced by novel requirements of Digital Libraries. These have been achieved through much exploratory work, primarily in the context of focused efforts devising specialized approaches to address particular aspects of Digital-Library functionality. For example, the ARTISTE project [2] from Europe’s Fifth Framework Programme focused on how to develop an integrated analysis and navigation environment for art images and analogous multimedia content, the COLLATE project [10] from the same Programme focused on how to deal with old film libraries, while the Alexandria Project [1] from NSF’s DLI-1 and DLI-2 Programs focused on geospatially-referenced multimedia material. For the most part, every effort so far has been distinct and, in some sense, isolated from the rest. Every project has started from scratch to build a system supporting the particular needs specified in the project’s description. Nevertheless, looking back at the individual achievements of all the projects, one may see clearly that there is substantial commonality among many of them; the bottom-up development of the field so far has provided enough ‘data points’ for patterns to emerge that can encapsulate all efforts.

Despite the young age of the field of Digital Libraries, it has made a long journey from its initial conception to the present state of the art and has reached a level of maturity that did not exist fifteen years ago. There is substantial knowledge and experience that have been accumulated. This warrants a process of self-declaration that will identify the principle ideas behind the field; it is time for a *Digital Library Manifesto* to set the ground rules for the field and lead to the development of reference documents that will capture the full spectrum of concepts that play a role in Digital Libraries.

---

As mentioned earlier, the nature of Digital Libraries is highly multidisciplinary. Naturally, this has created several conceptions of what a Digital Library is, each one influenced by the perspective of the primary discipline of the conceivers. In fact, Fox et al. [13] observe that the expression “Digital Library” evokes a different impression in each person, ranging from the simple computerisation of traditional libraries to a space in which people communicate, share, and produce new knowledge and knowledge products. For instance, the 1st Delos Brainstorming Workshop in San Cassiano, Italy, envisions a Digital Library as a system that enables any citizen to access all human knowledge, any time 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 [18]. An offspring of that activity concludes that Digital Libraries can become the universal knowledge repositories and communication conduits of the future, a common vehicle by which everyone will access, discuss, evaluate, and enhance information of all forms [19], [20]. Likewise, in his framework for Digital Library research, Soergel [28] starts from three very different perspectives that different people in the community have on Digital Libraries, i.e., as tools to serve research, scholarship, and education, as a means for accessing information, and as providing services primarily to individual users. He then enhances further each one and fuses them all together to obtain the main guiding principles for his vision of the field. On the other hand, Belkin [3] states that 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. Lesk [23] analyses and discusses the importance of the terms “Digital” and “Library” in the expression “Digital Library”, where the former term mainly implies existence of software for searching text, while the latter term refers to existing material that has been scanned for online access, and concludes that the research effort in the field are not usually associated with the users’ needs. Borgman [6] notices that at least two competing visions of the expression “Digital Library” exist: researchers view Digital Libraries as content collected on behalf of user communities, while practising librarians view Digital Libraries as institutions or services. Kuny and Cleveland [21] discuss four myths about Digital Libraries and attempt to bring them down: (i) the Internet is ‘The’ Digital Library; (ii) at some point there will be a single Digital Library or a single-window view of Digital Library collections; (iii) Digital Libraries are means to provide more equitable access to content from anywhere at any time; and (iv) Digital Libraries are cheaper instruments than physical libraries. They conclude that Digital Libraries impose reinvention of the role of librarians and library models.

In addition to such a variety of perspectives that may currently exist on what a Digital Library is, the concept has evolved quite substantially since the early idea of a system providing access to digitized books and other text documents. The DELOS Network of Excellence on Digital Libraries [11] now envisions a Digital Library as a tool at the centre of intellectual activity having no logical, conceptual, physical, temporal, or personal borders or barriers on information. It has moved from a content-centric system that simply organizes and provides access to particular collections of data and information, to a person-centric system that aims to provide interesting, novel, personalized experiences to users. Its main role has moved from static storage and retrieval of information to facilitation of communication, collaboration, and other forms of interaction among scientists, researchers, or the general public on themes that are pertinent to the information stored in the Digital Library. Finally, it has moved from handling mostly centrally-located text to synthesizing distributed multimedia document collections, sensor data, mobile information, and pervasive computing services.

This vision of Digital Libraries seems to resonate well with the concept of “Information Space” that has arisen from the field of Computer Supported Cooperative Work (CSCW).
Snowdon, Churchill, and Frecon [27] have developed future visions about “Connected Communities” and “Inhabited Information Spaces”, with the latter being closely related with the vision of Digital Libraries, in that ubiquitous information is a prerequisite for CSCW. In more detail, Inhabited Information Spaces are “spaces and places where people and digital data can meet in fruitful exchange, i.e., they are effective social workspaces where digital information can be created, explored, manipulated and exchanged”. Thus, “in Inhabited Information Spaces, both information and people who are using that information (viewing it, manipulating it) are represented. This supports collaborative action on objects, provides awareness of others’ ongoing activities, and offers a view of information in the context of its use”. Based on the above and according to the aforementioned DELOS vision of a Digital Library, the latter provides an Information Space that is populated by a user community and becomes an Inhabited Information Space through CSCW technology. The two fields complement each other nicely, in that one focuses on access and provision of relevant information while the other focuses on visualisation and sharing of information.

It becomes obvious that, as envisioned, “Digital Library” is a complex notion with several diverse aspects and cannot be captured by a simple definition. A comprehensive representation encapsulating all potential perspectives is required. This has 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. Having this broad scope, the Manifesto is followed by a set of separate reference documents, which stand individually but can also be seen as parts of a whole.

The ‘Digital Library Manifesto’ introduced here exploits the collective understanding that has been acquired on Digital Libraries by several previous efforts by European research groups active in the Digital Library field for many years, both within the DELOS Network of Excellence and outside, as well as by other groups around the world. Its role is one of a springboard for future work.

The following sections introduce the entities of discourse of the DL universe. They first introduce the relationships among three types of relevant “systems” in this area: Digital Library, Digital Library System, and Digital Library Management System. Then, they present the main concepts characterising the above, i.e., content, user, functionality, quality, policy, and architecture, and introduce the main roles that actors may play within a Digital Library, i.e., end-user, designer, administrator, and application developer. Finally, they describe the reference frameworks that are needed to clarify the DL universe at different levels of abstraction, i.e., the DL Reference model and the DL reference and concrete architectures.
2 The Digital Libraries Universe: A Three-Tier Framework

A Digital Library is a live organization that comes to exist through a series of development steps that bring together all necessary constituents. Figure 1 presents this process pictorially and indicates three distinct notions of “systems” that are developed along the way: Digital Library, Digital Library System, and Digital Library Management System. These also correspond to three different levels of conceptualization of the universe of Digital Libraries.

Unfortunately, these three system notions are often confused and used interchangeably in the literature. Nevertheless, as shown in Figure 1, they play a central and distinct role in the Digital Library development process. To clarify their differences and their individual characteristics, they are explicitly defined here as follows:

**Digital Library (DL)**

A (potentially virtual) organization 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 prescribed policies.

**Digital Library System (DLS)**

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

**Digital Library Management System (DLMS)**

A generic software system that provides the appropriate software infrastructure to both (i) produce and administer a Digital Library System that incorporates all functionality that is considered foundational for Digital Libraries and (ii) integrate additional software offering more refined, specialized, or advanced functionality.

A Digital Library Management System is what one may call “system software”. As in several other domains (e.g., operating systems, databases, user interfaces), such kernel software may be used as a foundation to produce Digital Library Systems. Depending on the philosophy it follows, a DLMS belongs to one of the following three categories:

- Extensible Digital Library System

---

**Figure 1. DL, DLS, and DLMS: A Three-Tier Framework**
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 [26] is a prototypical example of a system that is based on this
philosophy.

- **Digital Library System Warehouse**
  A collection of software components that cover all basic/foundational functionality
  required and a set of tools that can be used to combine these components in arbitrary
  ways (in Lego-like fashion) to create Digital Library Systems offering diverse sets of
  functionality. New software components can easily be incorporated into the
  Warehouse for subsequent combination with those already there. BRICKS [7] and
  DILIGENT [12] are two prototypical examples of systems that are based on this
  philosophy.

- **Digital Library System Generator**
  A highly parameterized software system that encapsulates templates for a broad range
  of functionality, which includes all basic/foundational as well as any advanced
  functionality that has been deemed appropriate. Through an initialization session, the
  appropriate parameters are set and tuned; at the end of that session, software is
  automatically generated that constitutes essentially a Digital Library System ready to
  be installed and deployed.

Note that, while the concept of Digital Library is intended to capture an abstract system that
consists of both physical and virtual components, the remaining two capture concrete software
systems. For every Digital Library, there is a unique Digital Library System in operation
(possibly consisting of many interconnected smaller DLSs in the most general case), whereas
all Digital Library Systems are based on a handful of Digital Library Management Systems.

It is also important to note that until now, most Digital Libraries have been created by
developing DLSs from scratch without any basis on a DLMS. The notion of DLMS has been
recently introduced by the DL community in order to overcome the many drawbacks arising
from the current DL development process [19][20]. When the digital library field reaches
maturity, one should expect that only a small number of industrial-strength Digital Library
Management Systems will be available on the market.

---

5 To the extent that it is helpful, one may draw an approximate analogy between the world of Digital Libraries and the world
of Databases. A DBMS (e.g., the DB2 or Oracle system) corresponds to a DLMS, offering general data management
services. A DBMS together with all application software running on top of it at an installation corresponds to a DLS. Finally,
a DL corresponds to a so-called “Information System” that consists of the above software, its data, and its users.
3 The Digital Libraries Universe: Main Concepts

Despite the great richness and diversity one finds among existing digital libraries, there are still a rather small number of concepts that are considered fundamental and essentially characterize all relevant systems and, in fact, the entire field. These concepts are identifiable in every digital library and should be the starting point for any researcher who wants to study and understand the field, for any system designer and developer who is about to embark in an effort to build a Digital Library, and for any content owner who wants to offer it externally. In this section, we identify these concepts and briefly discuss their roles.

Given that there are three kinds of “systems” in the digital library world, in principle each one may have, in addition to the common fundamental concepts, a number of other related concepts associated with it. Nevertheless, from the definitions of these systems in the previous section, one may easily derive that there is a natural inclusion relationship on the concepts that refer to these systems: the set of concepts used to define the Digital Library is completely included in that used for Digital Library System, which is in turn completely included in that characterising the Digital Library Management System. This inclusion relationship is depicted in Figure 2.

As a consequence of the above, we introduce below each of these concepts in reference to the highest level of a system (in the three-tier framework) for which it makes sense. All concepts, however, could also be described in reference to lower-level systems, focusing on the characteristics that are of primary concern there. More details can be found elsewhere [29].

The main concepts that are foundational in the Digital Library world are six. Five of them have been alluded to in the definition of Digital Libraries and are introduced in reference to them: Content, User, Functionality, Quality, and Policy; one of them has been alluded to in the definition of a Digital Library Systems and is introduced in reference to them: Architecture. All six concepts are shown together, influencing the Digital Library framework, in Figure 3 and are defined immediately after that.
Content
The Content concept represents the data and information that Digital Libraries handle and make available to their users. It is composed of a set of information objects organised in collections. Content is an umbrella concept that is used to aggregate all forms of information that a Digital Library may require to offer its services. For example, metadata play an important role in the information domain because they describe the contents of other information objects and express their structure.

User
The User concept represents the actors (whether human or not) entitled to interact with Digital Libraries. The aim of Digital Libraries is to connect such actors with information and to support them in consuming already available information and produce new information. As in the case of the previous concept, User is an umbrella concept that covers all notions related to the representation and management of actor entities within a Digital Library, i.e., the digital entities representing the actors, the rights that they have within the system, the profiles of the actors with characteristics that personalize the system’s behaviour or represent these actors in collaborations, and others.

Functionality
The Functionality concept represents the services that Digital Libraries offer to their users. In the most general case, the functionality expected from Digital Libraries is extremely rich. One may not expect, however, that all Digital Libraries must support the complete spectrum. Some core functionality represents the bare minimum that is required by all Digital Libraries, e.g., new information object registration, search, and browse. Beyond that, depending on the particular needs of its community of users and/or the information it is aimed to manage, each Digital Library may offer diverse additional functionality to serve those needs.

Quality
The Quality concept represents the parameters that can be used to characterize and evaluate the content and behaviour of Digital Libraries. It potentially includes several diverse
parameters associated with each class of content or functionality, possibly even with each distinct information object or service. 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 human experiments.

**Policy**

The Policy concept represents a set of rules that govern the interaction between users and Digital Libraries. Several mechanisms should be available to enable such interaction and the Digital Library should activate those mechanisms in the form prescribed by its policies. The policy concept includes all aspects of system security, digital rights management, system load control, and other similar issues.

**Architecture**

The Architecture concept (which refers to a Digital Library System) represents a mapping of the functionality and content offered by a DL onto hardware and software components. Having this as a main concept is important due to the following considerations: (i) Digital Libraries are often assumed to be among the most complex and advanced forms of information systems [14]; (ii) interoperability of Digital Libraries is one of the greatest challenges of the research community. Having a clear architectural understanding of Digital Library Systems offers ammunition in addressing both the above issues effectively.

In addition to the above primary Digital Library concepts, there is one more that is a derivative of them and is also very useful. That is the concept of Resource, which aggregates the concepts of Content, User, Functionality, and Policy, i.e., all concepts that refer to internal entities of a Digital Library and can be sensed by the external world. There are several aspects of all these concepts that are very similar or identical. Approached as instances of the more general concept of Resource, the Digital Library should treat them in the same fashion.

Having described the main concepts characterising the foundation of the digital library universe, we proceed in introducing four main roles of actors that interact with digital libraries at different levels and for different purposes.

---

6 This is an appropriate adaptation of the ‘Architecture’ definition from the Glossary of CMU’s Software Engineering Institute. [http://www.sei.cmu.edu/opensystems/glossary.html](http://www.sei.cmu.edu/opensystems/glossary.html)
4 The Digital Library Universe: The Main Roles of Actors

We envisage four different and complementary roles of actors interacting with digital libraries: DL End-Users, DL Designers, DL System Administrators, and DL Application Developers.

![Figure 4. The Main Roles of Actors versus the Three-tier Framework](image)

As depicted in Figure 4, each role is primarily associated with one of the three “systems” in the three-tier framework and, therefore, has the view that corresponds to that system. These roles are defined below.

4.1 DL End-Users

These 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. The state of the DL corresponds to the state of its resources, i.e., it consists of the collections of information objects managed by the DL, its set of authorized users, its functionality, and its set of policies. This state changes during the Digital Library lifetime according to the functionality activated by the users and their inputs. DL end-users may be further partitioned into Content Creator, Content Consumer, and Librarian.

4.2 DL Designers

These exploit their knowledge of the application semantic 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 former capture aspects of the DL functionality perceived by the end-users, e.g., the result set format, the query language, the user profile formats, and the document model. The latter capture third-party resources that are exploited by the specific DL, e.g., repositories of content, ontologies, classification schemas, authority files, and gazetteers. 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.

4.3 DL System Administrators

These select the software components necessary to create the Digital Library System needed to serve the required DL 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. Their task is to identify the architectural configuration that better fits the DLS in target ensuring the appropriate level of
quality. 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.

4.4 DL Application Developers

These develop the software components of DLMSs and DLSs, realizing the necessary functionality.

These four roles cover the whole spectrum of actors interacting with digital libraries. Their models of the digital library universe are linked together in a hierarchical fashion as depicted in Figure 5. This is a direct consequence of the above definitions since DL end-users act on the Digital Library, whereas the remaining three roles operate on the DLMS and consequently on the DLS and DL as well. This relationship ensures that roles designed to directly co-operate share a common vocabulary/knowledge. For instance, the DL end-user expresses its requirements in terms of the DL model and, subsequently, the DL designer understands these requirements and defines the DL accordingly.

![Figure 5. Hierarchy of Users' Views](image)
5 Reference Frameworks

The Digital Library Universe is quite complex and comprises multiple elements (see Figure 6). The representation of the details of these elements requires the introduction of several frameworks at different levels of abstraction:

- **Reference Model** – it is a set of inter-related concepts that collectively circumscribe and capture the essence of a distinct part of human knowledge, be it a whole scientific field or an individual idea or anything in between. It is a framework that helps in understanding of the basic elements of its concept(s) of concern. As stated in [24], “A Reference Model consists of a minimal set of unifying concepts, axioms and relationships within a particular problem domain, and is independent of specific standards, technologies, implementations, or other concrete details”. It can be established at various levels of abstraction, from very high-level and flexible to very concrete and precise. Although independent of them, it often leads to standards that have to be followed by all systems dealing with the particular parts of knowledge corresponding to it. Digital libraries need to obtain a corresponding reference model in order to consolidate the diversity of existing approaches into a cohesive and consistent whole and provide a common basis for further advancement.

- **Reference Architecture** – it is an architectural design pattern indicating an abstract solution in implementing the concepts and relationships identified in the Reference Model. Essentially, the reference architecture acts as a blueprint for the implementation of a DLS. There may be more than one reference architecture that addresses how to design digital libraries systems. Each of them proposes an optimal architectural pattern for a particular class of DLSs serving the needs of DLs characterised by similar goals, motivations and requirements. For example, we might have a reference architecture for DLSs supporting federated DLs, i.e. DLs built by federating local resources of multiple organizations, and another one for personal DLs, or for specialised applications.

- **Concrete Architecture** – it introduces additional elements making the reference architecture more concrete, e.g. replaces the mechanisms envisaged in the Reference Architecture with concrete standards and specifications. For example, it may specify that the run-time environment deployed on the hosting nodes is CORBA or the Web Services Application Framework, and that the Search functional component is implemented by four specific communicating Web Services.

The three frameworks above are shown in Figure 6 in relationship to the general digital library environment. At the top is the most abstract Reference Model, which guides the more specific Reference Architecture and Concrete Architecture further down. In turn, these should constrain the development of any actual Implementation of a system. The three reference frameworks are the outcome of an abstraction process that has taken into account all the natural Inputs, shown to the left, e.g., the systems already existing on the market, and has taken into consideration all existing Related Work, shown to the right, e.g., best practices and relevant research. When these frameworks are adopted and followed by the community, the resulting systems will be largely compatible with each other; the interoperability thus afforded will open up significant new horizons for the field.

---

7 This picture was inspired by the “Reference Model for Service Oriented Architecture” document [24].
Figure 6. The Digital Library Universe
6 Conclusions

The goal of ‘The Digital Library Manifesto’ has been to set the foundations and identify the entities of discourse within the universe of digital libraries. It has introduced the relationships among three kinds of relevant “systems” in this area: Digital Library, Digital Library System, and Digital Library Management System. It has presented the main concepts characterising the above, i.e., content, user, functionality, quality, policy, and architecture, and has identified the main roles that actors may play within a digital library, i.e., end-user, designer, administrator, and application developer. Finally, it has described the reference frameworks that are needed to clarify the digital library universe at different detailed levels of abstraction, i.e., the Reference Model and the Reference and Concrete Architectures.

‘The Digital Library Manifesto’ is accompanied currently by two other documents, which correspond precisely to the aforementioned reference frameworks. These documents come as a first attempt to fulfil the needs of the field in these areas. Clearly, diversity of needs among different digital libraries will continue to introduce new concepts that will require incorporation into the Reference Model and the Reference and Concrete Architectures. Hence, these documents should be considered as first versions of otherwise dynamic documents that will keep evolving, having the Manifesto as a firm foundation.

The ‘Digital Library Manifesto’ has been based on the experience and knowledge gained by many previous efforts that have gone on for the past fifteen years around Europe and the rest of the world. We hope it will serve as a basis for new advances in research and system development in the future.
Acknowledgements

The authors wish to acknowledge the considerable input received from the participants of the DELOS Reference Model Workshop held in Frascati, close to Rome, Italy, on June 2006. The comments, visions, and insights arising from the presentation of a first draft of this work and the ensuing discussions among all the experts that were present there have been extremely valuable in drafting the current version of the Manifesto. The workshop participants were the following: José Borbinha (DEI-IST-UTL), Martin Braschler (Zurich University of Applied Sciences Winterthur), Leonardo Candela (ISTI-CNR), Vittore Casarosa (ISTI-CNR), Tiziana Catarci (Università degli Studi di Roma “La Sapienza”), Donatella Castelli (ISTI-CNR), Stavros Christodoulakis (Technical University of Crete), Edward Fox (Virginia Tech), Norberth Fuhr (Universität Duisburg-Essen), Stefan Gradmann (Universität Hamburg), Yannis Ioannidis (University of Athens), Ariane Labat (EC), Mahendra Mahey (UKOLN), Patricia Manson (EC), Carlo Meghini (ISTI-CNR), Pasquale Pagano (ISTI-CNR), Andy Powell (UKOLN), Seamus Ross (University of Glasgow), Hans-Jörg Schek (ETH), Heiko Schuldt (University of Basel), MacKenzie Smith (MIT Libraries), Dagobert Soergel (University of Maryland), Costantino Thanos (ISTI-CNR), and Theo van Veen (National Library of the Netherlands).
References


The Digital Library Manifesto

Despite the large number of software tools named 'digital library systems', there is no agreement yet on what a Digital Library (DL) is and on the functionality that it must provide. Existing systems are heterogeneous in scope and range from digital object repositories to libraries, reference linking systems, archives and commercial systems that provide administration functions, to complete systems that integrate advanced digital library services and are mainly developed in research environments.

Research on DLS covers several different areas. It is often difficult to compare or combine the results achieved in these areas, since it is not always clear how they are related, and how they can mutually impact or constrain each other. This fragmentation of results hinders any effort to extend new research achievements into real-world systems.

These problems have a common origin: the lack of any agreement on the foundations of Digital Libraries.

DELOS has recently achieved the 'The Digital Library Manifesto' effort aimed to set the foundations and identify the commonalities and differences within the universe of DLS and a DL Reference Model, i.e., a conceptual framework representing the characteristics of the main 'systems' of the DL Universe. This model details the concepts related to the six main concepts of DL systems, i.e., content, user, functionality, quality, policy, and architecture, and it specifies the mutual relationships between these concepts and the constraints that hold among them. Both the manifesto and the reference model have been produced by exploring the experience that DELOS research groups have acquired over the years on the functionality and architecture of operational DLS.

began by a core group of DELOS members, the reference model activity now involves other DELOS and non-DELOS researchers, working from the perspectives of different domains. Laboratories with similar activities carried out by other international initiatives have been established in order to achieve a global and stable level of consensus on the model.