

Defining Digital Library

Armand Brahaj^{1,2}, Matthias Razum¹, and Julia Hoxha³

¹ FIZ Karlsruhe, 76344 Eggenstein-Leopoldshafen, Germany
{name.surname}@fiz-karlsruhe.de

² Humboldt-Universität zu Berlin, 10099 Berlin, Germany

³ Albanian Institute of Science, Tirana, Albania
julia.hoxha@ais.al

Abstract. This paper reflects on the range of the definitions of digital libraries demonstrating their extent. We analyze a number of definitions through a simplified intensional definition method, through which we exploit the nature of the definitions by analyzing their respective genera and attributes. The goal of this paper is to provide a synthesis of the works related to definitions of digital library, giving a fine-grained comparative approach on these definitions. We conclude that, although there are a large number of definitions, they are defined in overlapping families and attributes, and an inclusive definition is possible.

Keywords: Digital Library, Definition, Evaluation of Digital Libraries.

1 Introduction

The field of digital libraries (DL) has been for many years an avenue of extensive research and practical implementations. Despite the wide-ranging developments, the term “digital library” remains ambiguous and varies between different communities. Part of the ambiguity originates from the shared perceptions related to the concepts “digital” and “library”, but this has not spared the specialists who, approaching the problem from different backgrounds, bring their own conceptualization of a digital library.

The definition of DL is also an interesting topic because of its impact in the evaluation process. The connection between the evaluation and definition of the term has been advocated by Saracevic [1], who highlights the importance of DL definition in the evaluation process. Significant contributions in this regard have been introduced by the “Digital Library Reference Model” [2] proposed by DELOS. The topic is part of investigative research works such as [3] [4] [5]. These claims are also supported by Borgman [6] which points out that from a research perspective there are no attempts to introduce new definitions of DLs, but rather studies which synthesize already existing definitions. From the practice perspective she argues that there are additional needs such as evaluation metrics, which tend to introduce new definitions. The methods and metrics for the evaluation of DLs may vary according to whether these are viewed as institutions, information systems, new technologies, collections, or as services.

Motivated by the multitude of the various DL definitions, as well as their importance in further implementations and evaluation frameworks, this paper reflects on the range of the existing definitions demonstrating their respective extent. The work is based on an analysis of the nature of the genera and attributes across different

definitions. It is not the intent of the paper to introduce yet another digital library definition, but rather to provide a synthesis of the works related to the definitions of the digital library (in Section 3). Furthermore, we present an analysis of the definitions based on a fine-grained comparative approach (Section 4), concluding with the findings of the analysis and insights on the future use of the term.

2 Related Work

Discussions on the term “digital libraries” have been central to many scholarly communication activities since the notion came to existence. While the term dominates today in practice and research, the library community has used several different terms in the past [10]. Although with some resistance, the term “digital library” soon dominated and became widely accepted. Borgman [6] discussing the term concludes that its usage is problematic since it confuses the boundaries between electronic collections and institutions. She predicts that “*neither community is likely to surrender the term in favor of another. .. The failure to define the terms slows the development of theory, research and practice*” [6]. Careful approach in using the term has also been advised by Lynch [11]. Just as Borgman he regarded the term problematic considering the complex relation between the libraries as institutions and electronic collections. Harter [9] attempts to enclose the distance between traditional and DLs arguing that the latest should have the properties of a traditional library. Cleveland in [10] stresses the technical aspects of DLs as a step beyond the traditional library.

Significant contributions in the analysis of digital library have been done through the DELOS initiative in the Digital Library Manifesto [2]. The Manifesto led to the development of a reference document that captured key concepts involved in a DL. Manifesto’s analysis considers three types of relevant “systems” in this area: Digital Library as a virtual organization, Digital Library System as a specific interface system, and Digital Library Management System which is a software layer supporting administration. Their choice of separating these three concepts was a step ahead in avoiding confusion and the use of the notions interchangeably in literature.

In our quest to address the issues of the many different definitions we rely solely on the already existing definitions in attempt to show the extent of the concept. We do not analyze the definitions in the context of specific communities, but focus on the broad vision of digital libraries. Similar attempts are noted in [7] where he encourages a synthesis of definition.

3 Digital Library

3.1 Defining the Definition

Define: To state the precise meaning of a word or sense of a word [12]; To specify distinctly

A simple search for the definition of digital libraries in academic articles will yield more than a dozen results leading instinctively to the questions: Why are so many definitions for the same concept? Should the definitions be unique when related to a unique concept?!

We start our analysis by considering first the definition as a concept per se. The word *define*, as many other words used in the scientific terminology, derives from

Latin in the form *definire* and it is composed of the prefix *de-* and root *finire*. The prefix *de-* usually means "out of" or "away from." For example, *deplete* literally means "out of full" or "away from full". The meaning of *deplete* is close to *empty*. Just the same, *despera* means literally "out of hope".

The root *finire* indicates the extent of something, the limit as opposed to the word infinite. Deriving from the information above, the word *define* means away from the end/boundaries of something; putting an end to something as in enclosing; or as in showing its boundaries. Still, the meaning of the word *define* is not *confine*. While *confine* is more about keeping something restricted, the word *define* is about showing the potential *extent* of a concept. Following this reasoning, the many definitions that can be found on the term “digital library” are attempting to show the potential *extent* of DLs and also provide an overview of main components.

Definitions have also a well-organized form. A definition is a passage describing the meaning of a term, an object or a concept. Basically, in a definition we have [21]:

- a *definiendum* – the object of our definition,
- a *genus* – the family where the definition takes place
- and one or more *differentia* – distinguishable attributes

The above analysis of a definition will be the basis of the next section. If we take as an example the definition: “A *digital library* is an online system that stores media assets, and provides services for retrieving and presenting this content to humans or other online systems” then, the “*digital library*” is called the **definiendum**; *Online system* is the **genus**; *Stores media assets* and *provides services* are the **differentiae**.

In order to exploit the extent of DLs, we have used an analysis based on genus and differentia for a number of definitions found in scholar articles and institutional dictionaries. The method is also known as *intensional definition* [13] and it gives the meaning of a term by specifying all the properties required to derive that definition, i.e., the necessary and sufficient conditions for belonging to the set being defined.

We have applied this analysis to a list of prominent definitions on digital libraries. Even though we claim that this paper provides an exhaustive coverage of works that contain definitions of digital library, we are aware that potentially there can be other definitions which could be appended to our analysis. Some of the extracted genera and the differentiae are illustrated in Table 1 together with the respective definitions. An extensive datasets and definitions used in our analysis can be found at [8].

Table 1. Analysis of Digital Library definitions, partial dataset

Digital Libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities[14]	Genus	Organization
	Differentia	Provide resources and specialized staff to: - select collections of digital work - structure collections of digital work - offer intellectual access to collections of digital work - interpret collections of digital work - distribute collections of digital work - preserve the integrity of digital work - ensure persistence over time of digital work collections

An informal definition of a digital library is a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network[15]	Genus	Managed collection
	Differentia	Associated services Information is stored in digital format Information accessible over network
Digital library is "a focused collection of digital objects, including text, video, and audio, along with methods for access and retrieval, and for selection, organization, and maintenance of the collection[16]	Genus	Focused collection
	Differentia	Methods for access and retrieval of collection
Digital libraries viewed as systems providing a community of users with coherent access to a large, organized repository of information and knowledge [17]	Genus	Systems
	Differentia	Provide community with access to large, organized information
A digital library is a distributed technology environment which dramatically reduces barriers to the creation, dissemination, manipulation, storage, integration, and reuse of information by individuals and groups. [18]	Genus	Distributed Environment
	Differentia	Reduces barriers to information: creation, manipulation, storage, reuse
A digital library is an integrated set of services for capturing, cataloging, storing, searching, protecting, and retrieving information. [19]	Genus	A group of Services
	Differentia	These services enable: capturing, cataloging, searching, protecting, retrieving, information
A "digital library" is fundamentally a resource that reconstructs the intellectual substance and services of a traditional library in digital form. [7]	Genus	Resource
	Differentia	Reconstruct the intellectual substance Reconstructs the services of a traditional library in digital form
A possibly 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 codified policies [2]	Genus	Virtual organization
	Differentia	Collects, manages and preserves digital content; Offers specialized content functionality to user communities and according to codified policies

3.2 Synthesis of the Definitions

We have considered a number of definitions on digital library and applied an analysis of the definition to each of them. The result is an extraction of genera and differentia in each definition. After analyzing the genera extracted from the definitions, we realized that the genera can be grouped under four main concepts: Collection, Service, Organization, and System

As illustrated in Table 2, most of the definitions bind digital libraries to Collections. It is also interesting to find out that most of the definitions that relate digital library to Collections highlight specifications of collections, such as focused collection, managed collection or organized collection. The requisite to explicitly specify collections shows a bond of the concept with specific perceptions related to usability

Table 2. Genera of digital libraries grouped in four groups

Collections	Services	Organization	Systems
<ul style="list-style-type: none"> - Organized collections - Managed collection - Focused collection - Electronic resources - Collection of collections - Collection of information objects - Collection of documents in electronic format - Resource 	<ul style="list-style-type: none"> - Library services - Dynamic federated structures - Information storage - Retrieval systems - Distributed environment - Collection of services - Group of services 	<ul style="list-style-type: none"> - Organization - Operational organization - Socio-technical systems - Virtual organization 	<ul style="list-style-type: none"> - Systems - Tools - Electronic resources - Database on hypertext environment - Environment - Library - Socio-technical systems - Networks of technology

and added value. The terminology still varies in different definitions, sometime focusing on the extension of library services, sometime on interoperability (dynamic federated structures), and in other cases on information storage and information retrieval systems. Surprisingly, few definitions regard digital libraries as *Organizations*, although many properties tend to push toward this family. A more abstract genus is found in defining DLs as *Systems* leaving space for technical interpretations.

Just as the genera of the DL overlap, so do the differentia found in different definitions. Our analysis on the attributes provides a first attempt towards generalization and categorization. It is interesting to notice the occurrence of these attributes in the list of candidate properties which are needed in the evaluation of digital libraries as proposed by Saracevic [1]. Many of the attributes can also be seen as mapped inside the *main concepts* of the DELOS's Digital Library Universe. Their analysis is an interesting direction for future research.

Recalling the meaning of the term definition itself (Section 4.1) as the potential extent of a concept, as well as the observed overlap between the genera and differentia found among the existing definitions of digital library, we strongly argue that this collection of genera and differentia constitutes a basis for an inclusive definition of DL.

4 Conclusions

The use of the term 'digital library' is very broad. Digital libraries have evolved dynamically over the past two decades, and such has the use of the term. Digital libraries are no longer theoretical constructs, but reality. They are driven by user groups or communities. Each of them focuses on their specific usage scenarios, leading to the existing variety of definitions. In most of the definitions evaluated in this work, community- and scenario-specific requirements are not expressed as genera, but rather as attributes found in differentia. Separating the manifold of attributes from the overarching concepts of 'collection', 'service', 'organization', and 'service' allows us to provide a framework for existing and future definitions of the term 'digital library'. Every definition should reflect those concepts. This does not necessarily mean that all four concepts must be part of the definition. Depending on the concrete usage scenario, one or more genera might be left out intentionally. However, given the prominence of these genera throughout all evaluated existing definitions, they provide a valid guideline for future work.

For this paper, we did not analyze the attributes found in the differentia of the evaluated definitions in greater detail. A rather cursory grouping produced the following, non-definitive list: Intellectual Access, Service Management, Information Structuring, Collection Management, Digital Accessibility, Permanent Access, Supplement to conventional library collections, Economic Perspective, Technical Capabilities. A more systematic and accurate grouping is an avenue for future work.

Although we initially started this work with the conception that the term ‘digital library’ is ill-defined and maybe even misleading, the results of the analysis showed that it is based on four abstract concepts, that are then concretized with community- and scenario-specific attributes.

References

1. Saracevic, T.: Digital library evaluation: Toward evolution of concepts. *Library Trends* 49(2), 350–369 (2000)
2. Candela, L., et al.: The DELOS Digital Library Reference model. *Foundations for digital Libraries (Version 0.98)* (2008)
3. Fuhr, N., et al.: Evaluation of digital libraries. *International Journal on Digital Libraries* 8(1), 21–38 (2007)
4. Xie, H.I.: Users’ evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment. *Information Processing & Management* 44(3), 1346–1373 (2008)
5. Khoo, M., MacDonald, C.: An Organizational Model for Digital Library Evaluation. In: Gradmann, S., Borri, F., Meghini, C., Schuldt, H. (eds.) *TPDL 2011*. LNCS, vol. 6966, pp. 329–340. Springer, Heidelberg (2011)
6. Borgman, C.L.: What are digital libraries? Competing visions. *Information Processing Management* 35, 227–243 (1999)
7. Seadle, M., Greifeneder, E.: Defining a digital library. *Library Hi Tech* (2007)
8. Brahaj, A., Razum, M., Hoxha, J.: Dataset of definitions related to Defining Digital Libraries. Figshare (2013), <http://dx.doi.org/10.6084/m9.figshare.707337>
9. Harter, S.P.: What is a digital library? Definitions, content, and issues. In: *International Conference on Digital Libraries and Information Services for the 21st Century* (1996)
10. Cleveland, G.: Digital Libraries: Definitions, Issues and Challenges. In: *UDT* (1998)
11. Lynch, C.A.: Accessibility and integrity of networked information collections. *Congress of the United States, Office of Technology Assessment, Washington, DC* (1993)
12. American Heritage, *American Heritage® Dictionary of the English Language*, 4th edn. (2003), <http://www.thefreedictionary.com/define> (accessed 2013)
13. Cook, R.T.: Intensional Definition. In: *A Dictionary of Philosophical Logic*, p. 155. Edinburgh University Press (2009)
14. DLF (1998), <http://old.diglib.org/about/dldefinition.htm> (accessed 2013)
15. Arms, W.: *Digital Libraries*. M.I.T. Press (2000)
16. Witten, I.H., et al.: *How to build a digital library*. Morgan Kaufmann (2009)
17. Lynch, C., Garcia-Molina, H.: Interoperability, scaling, and the digital libraries research agenda. In: *Iita Digital Libraries Workshop* (1995)
18. Fox, E.A.: *Source Book on Digital Libraries*. Virginia Tech, Virginia (1993)
19. WTEC Hyper-Librarian, *WTEC Principles of Digital Libraries* (1999), http://www.wtec.org/loyola/digilibs/04_02.htm (accessed 2013)
20. Ojha, R.C., Aryal, S.: *Digital Libraries: Challenges and Opportunites*. Infolib (2010)
21. Copi, I.M., Cohen, C.: *Introduction to Logic*, 9th ed (1994)