

# (Formal) Models for Systems, Infrastructures, Communities, and Cultures

Nicola Ferro

Department of Information Engineering, University of Padua, Italy  
ferro@dei.unipd.it

**Abstract.** This paper highlight the role (formal) model for digital libraries can play to gap the bridge between different communities and cultures, such as libraries and archives, in order to enable interoperability among systems and infrastructures.

**Keywords:** DELOS Reference Model, 5S Model, Libraries, Archives

## 1 Digital Libraries Models

Since the field of digital libraries has come to light in the early nineties of the past century, a lot of improvements and a dramatic change in the viewpoint has happened. In the beginning, digital libraries were almost monolithic systems, each one built for a specific kind of information resources – e.g. text, images, or videos – and with very specialised functionalities developed ad-hoc for those contents. This approach caused a flourishing of systems where the very same functionalities, e.g. user management or repositories, were developed and re-developed from scratch many times, causing them to be different and often incompatible one with the other.

With the passing of time and by exploiting the previous research results and achievements, a more mature way of facing the design and development of digital libraries has taken place. Digital libraries moved from being monolithic systems to being component and service-base systems, where easily configurable and deployable services can be plugged together and re-used in order to create a digital library. Moreover, digital libraries started to be seen as more and more user-centered systems, where the original content management task is partnered with new communication and cooperation tasks, so that digital libraries become “*a common vehicle by which everyone will access, discuss, evaluate, and enhance information of all forms*” (Ioannidis et al., 2005).

In this evolving scenario, the design and development of effective services which foster the cooperation among users and the integration of heterogeneous information resources become a key factor which needs to be pursued by researchers and developers. A relevant example of this kind of new services are annotations, i.e. providing users or groups of users with the possibility of adding personal annotations on the managed information resources, even crossing the boundaries of the single digital library (Agosti et al.; 2013; Agosti and Ferro, 2008).

In this context, building foundations and a formal theory for digital libraries is a longstanding issue in the field, dating back to the mid-1960s (Licklider, 1965), and this challenge has been accepted only very recently, for example, by the 5S model (Gonçalves et al., 2004) and the DELOS Reference model (Candela et al., 2007).

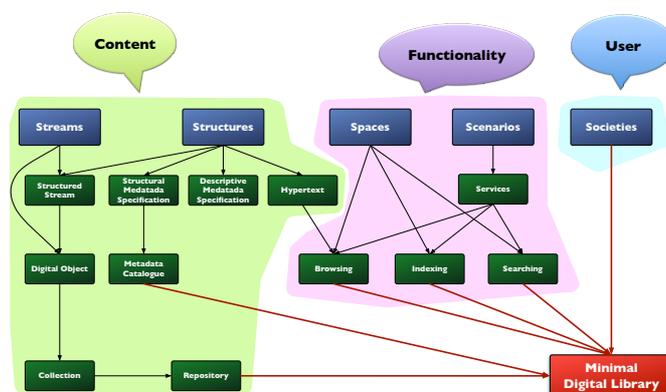
## 1.1 The DELOS Reference Model

The DELOS Reference Model lays the foundations of digital libraries and defines what are the constituent entities and stakeholders of the digital library universe as well as the relationships among them; in particular, the reference model provides a clear picture of what a digital library is and on what concepts and functionalities we can leverage in order to promote co-operation and interoperability (Candela et al., 2007). The DELOS Reference Model approaches the problem of modelling the digital library universe by highlighting six domains or main concepts: *content* is the data and information that digital libraries handle and make available to their users; *user* is the actors (whether human or not) entitled to interact with digital libraries; *functionality* is the services that digital libraries offer to their users; *quality* is the parameters that can be used to characterize and evaluate the content and behaviour of digital libraries; *policy* is a set of rules that govern the interaction between users and digital libraries; *architecture* is a mapping of the functionality and content offered by a digital library onto hardware and software components.

Moreover, the DELOS Reference Model distinguishes among three different “systems” which constitute the digital library universe and rely on the six domains introduced above for their definition: *Digital Library (DL)* is an organisation, which might be virtual, that comprehensively collects, manages and preserves for the long term rich *digital content*, and offers to its *user* communities specialised *functionality* on that content, of measurable *quality* and according to codified *policies*; *Digital Library System (DLS)* is a software system that is based on a defined (possibly distributed) *architecture* and provides all functionality required by a particular Digital Library. Users interact with a Digital Library through the corresponding Digital Library System; *Digital Library Management System (DLMS)* is a generic software system that provides the appropriate software infrastructure both (i) to produce and administer a Digital Library System incorporating the suite of functionality considered fundamental for Digital Libraries and (ii) to integrate additional software offering more refined, specialised or advanced functionality.

## 1.2 The 5S Model

The Streams, Structures, Spaces, Scenarios, Societies (5S) (Gonçalves et al., 2004) is a formal model for digital libraries based on the following abstractions: *Streams* are sequences of elements of an arbitrary type (e.g. bits, characters, images) and thus they can model both static and dynamic content; *Structures* are the way through which parts of a whole are organised. In particular, they can be used to represent hypertexts and structured information objects, taxonomies, system connections and user relationships; *Spaces* are sets of objects together with operations on those objects conforming to certain constraints; *Scenarios* are sequences of events that may have parameters, and events represent state transitions; *Societies* are sets of entities and relationships. The entities may be humans or software and hardware components, which either use or support digital library services.



**Figure 1: Main definitions of the 5S model and their relationships with the domains of the DELOS Reference Model.**

As shown in Figure 1, from the five abstractions of streams, structures, spaces, scenarios, and societies, a series of concepts are derived, which are then used to define what a digital library is. Indeed, in accordance with this framework, a minimal digital library is defined a constituted by: a *repository*, that is a service encapsulating a family of collections and specific services to manipulate the collections; a set of *metadata catalogues* for all the collections in the repository; a set of *services* containing, at least, services for indexing, searching and browsing; and, a society whose information needs have to be satisfied. As you can note from Figure 1, only three out of the six domains of the DELOS Reference Model are taken into consideration in the 5S model, namely the Content, Functionality, and User domains; the other three – Quality, Policy, and Architecture – are not dealt with but are left to additional models that can be built starting from the 5S model.

## 2 Bridging between Libraries and Archives

In the context of Libraries, Archives, and Museums (LAM) unifying a variety of organizational settings and providing more integrated access to their contents is an aspect of utmost importance. Indeed, LAM collect, manage and share digital contents; although the type of materials may differ and professional practices vary, LAM share an overlapping set of functions. Fulfilling these functions in “collaboration rather than isolation creates a win-win for users and institutions” (Zorich et al., 2008). Although the convergence between libraries, archives and museums has been a topic of much discussion in the digital library community, the emerging similarities between these three types of cultural heritage institutions are not yet evident in the proposed formal models, developed systems, and education of professionals (Trant, 2009; Timms and Fall, 2009).

Archives are a fundamental constituent of our cultural heritage and digital libraries are the natural choice for managing and providing access to their assets. Unfortunately, there have been almost no formal models for archives and this has prevented them from being fully integrated in digital library communities, methodologies and technologies. We think that the archival domain deserves a formal theory as well and that

this theory has to be reconciled with the more general theories for digital libraries in order to provide archives with the full breadth of methodologies and technologies which have been developed over the last two decades in the digital library field.

We proposed the NESTOR formal model (Ferro and Silvello, 2013) to settle a common ground for dealing with hierarchies open to existing models, solutions and technologies. The set data models composing NESTOR are well-suited for archival practice; indeed, the idea of “set” shapes the concept of archival division which is a “container” comprising distinct elements that have some properties in common. If we consider the Chinese boxes metaphor, a hierarchy is composed of a sequence of boxes contained one inside the other; if we look at an archive from the physical point-of-view, we can see that it resembles the Chinese boxes structure as there are boxes, folders, sheets, etc. contained one inside the other. Nested sets are closer to this view of reality than trees are. Indeed, although archival practice commonly considers archives as trees, a tree is actually a higher level abstraction than the nested sets as it only focuses on structural relationships. Indeed, NESTOR comprises both the structure and the content of the archive, where the inclusion relationships represent the structure and the elements belonging to the sets represent the content.

Then, we extended the 5S model to introduce the notion of digital archive as a specific case of digital library complying with the NESTOR archival constraints. This, in turn, will open up the possibility to further extend the 5S model. Indeed, according to this model, a minimal digital library has to offer indexing, searching and browsing services (Gonçalves et al., 2004). The formal definition of the query and update operations in NESTOR will thus allow us to precisely describe what these services are in the case of digital archives.

## References

- Agosti, M., Conlan, O., Ferro, N., Hampson, C., and Munnely, G. (2013). Interacting with Digital Cultural Heritage Collections via Annotations: The CULTURA Approach. In *Proc. 13th ACM Symposium on Document Engineering (DocEng 2013)*, pages 13-22. ACM Press, New York, USA.
- Agosti, M. and Ferro, N. (2008). A Formal Model of Annotations of Digital Content. *ACM Transactions on Information Systems (TOIS)*, 26(1):3:1–3:57.
- Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobreva, M., Katifori, V., and Schuldt, H. (2007). The DELOS Digital Library Reference Model. Foundations for Digital Libraries. ISTI-CNR at Gruppo ALI, Pisa, Italy.
- Ferro, N. and Silvello, G. (2013). NESTOR: A Formal Model for Digital Archives. *Information Processing & Management*, 49(6):1206-1240.
- Gonçalves, M. A., Fox, E. A., Watson, L. T., and Kipp, N. A. (2004). Streams, Structures, Spaces, Scenarios, Societies (5S): A Formal Model for Digital Libraries. *ACM Transactions on Information Systems (TOIS)*, 22(2):270–312.
- Ioannidis, Y., Maier, D., Abiteboul, S., Buneman, P., Davidson, S., Fox, E. A., Halevy, A., Knoblock, C., Rabbitt, F., Schek, H.-J., and Weikum, G. (2005). Digital library information-technology infrastructures. *International Journal on Digital Libraries*, 5(4):266–274.
- Licklider, J. C. R. (1965). *Libraries of the future*. Cambridge, Massachusetts, USA: The MIT Press.
- Timms, K., & Fall (2009). New partnerships for old sibling rivals: The development of integrated access systems for the holdings of archives, libraries, and museums. *Archivaria*, 68, 67–96.
- Trant, J. (2009). Emerging convergence? Thoughts on museums, archives, libraries, and professional training. *Museum Management and Curatorship*, 24(4), 369–387.
- Zorich, D. M., Waibel, G., Erway, R., & Zorich (2008). Beyond the silos of the LAMs: Collaboration among libraries, archives and museums. Tech. rep., OCLC Programs and Research, Dublin, Ohio, USA.