Virginia Tech Computer Science Technical Report - March, 2010

Extending the 5S Framework of Digital Libraries to support Complex Objects, Superimposed Information, and Content-Based Image Retrieval Services

Uma Murthy, Nádia P. Kozievitch, Jonathan Leidig, Ricardo Torres, Seungwon Yang, Marcos Gonçalves, Lois Delcambre, David Archer, and Edward A. Fox

the date of receipt and acceptance should be inserted later

Abstract Advanced services in digital libraries (DLs) have been developed and widely used to address the required capabilities of an assortment of systems as DLs expand into diverse application domains. These systems may require support for images (e.g., Content-Based Image Retrieval), Complex (information) Objects, and use of content at fine grain (e.g., Superimposed Information). Due to the lack of consensus on precise theoretical definitions for those services, implementation efforts often involve ad hoc development, leading to duplication and interoperability problems. This article presents a methodology to address those problems by extending a precisely specified minimal digital library (in the 5S framework) with formal definitions of aforementioned services. The theoretical extensions of digital library functionality presented here are reinforced with practical case studies as well as scenarios for the individual and integrative use of services to balance theory and practice. This methodology has implications that other advanced services can be continuously integrated into our current extended framework whenever they are identified. The theoretical definitions and case study we present may impact future development efforts and a wide range of digital library researchers, designers, and developers.

Uma Murthy · Jonathan Leidig · Seungwon Yang · Edward A. Fox
Department of Computer Science, M/C 0106,
Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
E-mail: {umurthy,leidig,seungwon,fox}@vt.edu

Nádia P. Kozievitch · Ricardo da S. Torres
Institute of Computing, University of Campinas,
Av. Albert Einstein, 1251, CEP 13084-851, Campinas, SP, Brazil
E-mail: {nadiapk,rtorres}@ic.unicamp.br

Marcos André Gonçalves
Department of Computer Science,
Federal University of Minas Gerais, Belo Horizonte, Brazil
E-mail: mgoncalv@dcc.ufmg.edu

Lois Delcambre and David Archer
Portland State University, Portland, OR 97207, USA
E-mail: {lmd,darcher}@cs.pdx.edu

1 Introduction

Users involved in creation of, management of, and access to media of all types are often concerned about improving productivity. At the same time, the volume and assortment of media content to be considered in such tasks continues to grow exponentially. As a result, users increasingly turn to advanced integrated information systems in order to assist them in their work. Digital libraries (DLs) are widely used for such tasks. However, some DLs provide only simple services, e.g. metadata text searching or full-text indexing. Few DLs provide services in support of newer, more complex media types like images, multimedia objects, subdocuments within other documents, or annotations. In addition, there is little evidence of common vision among digital library architects for providing commonality or uniformity of services for these media among DLs. We believe the DL community may benefit from formal definitions of existing and proposed digital libraries along with the services they provide, so that future DL architects may provide more uniform services, and may re-use existing service modules in new library architectures. In this paper, we address formal definitions and descriptions of desired functionality for DLs in three areas: 1) complex objects (i.e., digital objects that consist of two or more other digital objects); 2) superimposed information and services which involve fine-grain contextualized information found in parts of documents; and 3) content-based image retrieval (CBIR) services dealing with retrieval of images, considering content features such as color, shape, and texture.

1.1 Reference Models

Previous efforts in defining digital libraries have progressed towards a digital library reference model. Two such efforts have led to the 5S framework [28, 31] and the DELOS Reference Model [20, 19]. Nevertheless, there is no universally accepted reference model that defines all aspects of DLs in a precise and rigorous fashion.

1.1.1 5S Framework

The 5S framework aims to provide an underlying foundation for the definition of digital libraries [31, 30]. The unified formal theory specifies Streams, Structures, Spaces, Scenarios, and Societies. In turn, these can be employed to describe other key DL concepts, such as digital objects, metadata, collections, and services. *Streams* are sequences of elements of particular types (e.g., characters, pixels, bits, etc.). Streams are used to model static and dynamic content, including textual material and multimedia content. *Structures* specify the way in which parts of DLs are organized. Structures are used to represent hypertexts, system organization, and containment. *Spaces* are sets of objects and their operations. Spaces define the logical and presentational views of many DL components (e.g., probability and vector spaces). *Scenarios* are sequences of events along with a number of parameters. Events may represent changes in computational states through specific parameter values. Behaviors of DL services are described using scenarios. *Societies* are used to describe entities with their relationships to other entities. Societies may include human users and software entities that play a defined role in the digital library's operation.

These five abstractions are useful in providing a foundation for defining and relating digital library concepts. As an example, a digital object may be defined in terms of its *structured* storage *stream* and *structured* metadata specification. The set of 5S descriptions for a digital library may be encompassed in XML 5SL representations. The 5SL representations may be used to generate and install an implementation of the described digital library [29].

1.1.2 5S Minimal Digital Library

5S framework efforts to date have focused on defining the minimal set of features belonging to a digital library. The minimal digital library is defined as a quadruple (Repository, Metadata Catalogue, Services, Society) containing the core digital library components. These features include the basic set of structured content and elementary services provided to end-users. In a minimal digital library, the Structures component is missing; digital objects are represented through one or more streams and have an associated metadata record with a simple structure. There are not other structures in a minimal digital library. The basic digital library services include indexing, searching, browsing, and visualizations [68]. Minimal digital libraries have been defined with minor context-specific additions to produce existing libraries such as the archaeological ETANA-DL [66]. Further refinement of the 5S framework aims to extend the functionality beyond that of the minimal digital library.

1.1.3 DELOS Reference Model

The DELOS Reference Model [19, 20] is a similar effort in digital library foundations that places less emphasis on using abstract concepts to represent system components. The DELOS Reference Model consists of three tiers: digital library, digital library system, and digital library management system. There are six main concepts: content, user, functionality, quality, policy, and architecture. These concepts are used to directly describe digital library aspects by informal methods. The DELOS model focuses on identifying the main concepts and relationships encompassing the entire digital library as opposed to defining individual digital library aspects in terms of abstract entities. The formal definitions that are present in the 5S framework have been considered a future step for the DELOS reference model.

1.2 Applications and Scenarios

Many digital library implementations will require additional services beyond that of the minimal digital library. For example, e-Science or cyberinfrastructure applications typically require large datasets and high-performance computing (HPC) resources. Management systems for e-Science applications must be able to accept HPC input parameters and process large amounts of data. Depending on the application, some of the lower value datasets that are generated may be discarded. Support is likely required for storing and managing input parameters, underlying datasets and models, raw computational outputs, analyses, and publications. An example is the set of experiments and findings [15] derived from a computational epidemiology simulation system [14].

Additional functionality is needed for biological research. The identification of fish species is an example of a desired capability of a biological research system [47]. These applications require the management of images, text, and annotations. Users may search with keyword descriptions or hope to match personal fish images with identified fish in the collection. Digital libraries in this context may be required to provide specialized support such as image processing algorithms for fish contours.

See Table 1 for a listing of scenarios for our current efforts to extend the 5S framework to meet a range of functionality requirements. Note that these scenario examples detail

Table 1 Examples of individual and integrative services of a DL

Definition	Systems	Descriptive Example
CBIR	[26, 32, 78]	A Veterinary student attempts to find all the images that are similar to the one that he specifies (Scenario No.1 in Appendix)
Complex objects	[61, 62]	A parasite researcher deals with potentially heterogeneous data and metadata as a unified group (Scenario No.2 in Appendix)
Superimposed information	Xanadu [63], Flickr [5] notes, combinFormation [40], Superimposed applications [48, 51, 52]	A Computer Science professor works with pieces of information to prepare her lecture (Scenario No.3 in Appendix)
Integrated Service	[37, 14, 15]	An Epidemic simulation researcher stores results along with related information (Scenario No.4 in Appendix)
Integrated Service	[56, 58, 59]	A student in Fisheries takes a test on fish species identification (Scenario No.5 in Appendix)

the interaction between the user and system. In the 5S framework, scenarios refer to the behavior of the system. Each of the first three scenarios are examples of use of specific functionality that may result from one new feature among: complex objects, superimposed information, and CBIR. The last two scenarios illustrate the use of integrated functionalities provided by combining two or more of the extensions. The extensions mentioned in this paper cover a subset of practical scenarios and services as needed by users today. Our plan is to develop a series of incremental extensions, each precisely specified and adding key services for important scenarios, so eventually all that is covered in the DELOS Reference Model [19, 20], and more, is incorporated.

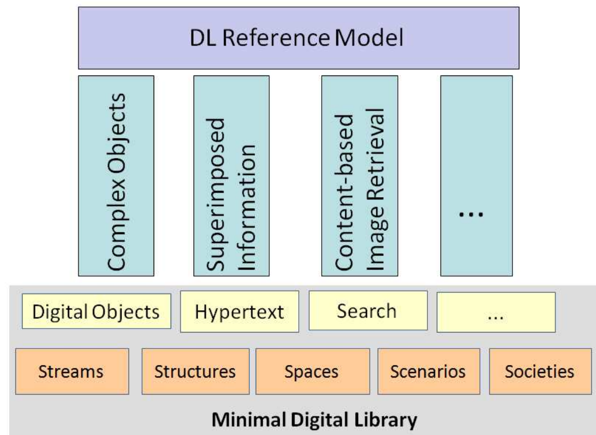

Fig. 1 Extensions towards a reference model.

1.3 5S Extensions

The definitions found in this paper extend the existing 5S framework [58], working towards comprehensive coverage for a reference model as shown in Figure 1. Services for content based image retrieval (CBIR), complex objects (CO), and superimposed information (SI) are commonly required in digital library systems (as also shown in Table 1). The prevalence of these three digital library aspects led us to extend the 5S framework to define each of the three topics. CO and SI extend the notion of digital objects, as described in the 5S framework. Working with SI involves referencing fine-grain information in documents. This link relationship extends the idea of hypertext. Finally, CBIR may be considered as advanced image search.

The inclusion of more topical extensions will work towards the coverage of concepts in the DELOS Reference Model. The main contribution of this paper is that it demonstrates how the 5S framework may be extended to provide support for complex objects, content-based image retrieval services, and superimposed information and services. These constructs may be reused in future 5S descriptions and extended in further 5S supplementation efforts.

The rest of this document is organized as follows. Section 2 contains an overview and description of related work in complex objects, superimposed information, and content-based image retrieval. The preliminary work on formalizing complex objects using the 5S framework is described in Section 3. Section 4 contains the 5S extensions for superimposed information. The 5S extensions for content-based image retrieval are presented in Section 5. Lastly, we present a case study and our conclusions in Sections 6 and 7 respectively.

2 Related Work

2.1 Complex Objects

Several complex object (CO) formats arise from different communities [61, 62]. In scientific computing, standards arise, such as Network Common Data Form (NetCDF), Hierarchical Data Format (HDF), and Extensible File System (ELFS). HDF and NetCDF, for example, are used in multi-dimensional storage and retrieval, while ELFS is an approach to address the issue of high performance I/O by treating files as typed objects.

COs often are found in persistent database stores. They may be represented using standards like MPEG-21 [18] or

METS [25]. Other technologies have been proposed, as standard Moving Picture Experts Group (MPEG), multimedia framework MPEG-21 and digital object formats as Moving Picture Experts Group - 21 Digital Item Declaration Language (MPEG-21 DIDL) and Metadata Encoding and Transmission Standard (METS).

Even though there are a number of standards aiding in the management of COs, there is still incompatibility, motivating solutions for integration and interoperability. As each standard is specialized for a particular domain, it is hard to interoperate across contexts.. Yet, it is to match some of them, as proposed in [24], in their comparative study of IMS Content Package (IMS CP) and Reusable Asset Specification (RAS).

New standards have emerged, like SQL Multimedia and Application Packages (SQL/MM) [50]. These were defined to describe storage and manipulation support for complex objects. A number of candidate multimedia domains were suggested, including full-text data, spatial data, image data, and others.

The Open Archival Information System (OAIS) [69] is an International Organization for Standardization (ISO) reference model, with a particular focus on digital information, both as the primary form of information held and as supporting information for both digitally and physically archived materials. The objects are categorized by their content and function in the operation of an OAIS into Content Information objects, Preservation Description Information objects, Packaging Information objects, and Descriptive Information objects. The Content Information is the set of information that is the original target of preservation by the OAIS. In addition to Content Information, the Archival Information must include information that will allow the understanding of the Content Information over an indefinite period of time (Preservation Description Information objects). The Packaging Information is that information which, either actually or logically, binds or relates the components of the package into an identifiable entity on specific media. And finally, in addition to preserving information, the OAIS must provide adequate features to allow Consumers to locate information of potential interest, analyze that information, and order desired information (Descriptive Information objects).

The Open Archives Initiative (OAI) [42] is a framework for archives (institutional repositories) containing digital content (digital libraries). The OAI technical infrastructure, specified in the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) [70], defines a mechanism for data providers to expose their metadata. This protocol mandates that individual archives map their metadata to the Dublin Core, a simple and common metadata set for this purpose.

OAI later launched the Object Reuse and Exchange (OAI-ORE) [46] project which defines standards for the description and exchange of aggregations of Web resources, and is developing interoperable, and machine-read-able mechanisms to express compound object information on the web. OAI-ORE makes it possible to reconstruct the logical boundaries of compound objects, the relationships among their internal components, and their relationships to the other resources. The information is encapsulated with named graphs: a set of RDF assertions identified by a URI. A named graph can be described by a resource map. OAI-ORE uses the web architecture [44], essentially consisting of:

- URIs for identifying objects;
- resources, which are items of interest;
- standard protocols, such as HTTP, that enable the access;
- links via URI references;
- named graphs for encapsulating information into a compound object.

METS [25], addresses packaging to collect digital resource metadata for submission to the repository. It is a Digital Library Federation initiative. A METS document consists of the following sections: header, descriptive metadata, administrative metadata, file section, structural map, structural links, and behavior.

METS uses a structural map to outline a hierarchical structure for the digital library object, where file elements may be grouped within fileGrp elements, to provide for subdividing the files by object version. A $\langle fileGrp \rangle$ structure is used to comprise a single electronic version of the digital library object. $\langle FContent \rangle$ was created to embed the actual contents of the file within the METS document, but it is rarely used. METS provides an XML Schema designed for the purpose of:

- Creating XML document instances that express the hierarchical structure of digital library objects.
- Recording the names and locations of the files that comprise those objects.
- Recording associated metadata.

METS can, therefore, be used as a tool for modeling real world objects, such as particular document types.

SCORM [45] is a compilation of technical specifications to enable interoperability, accessibility and reusability of web-based learning content. With a Content Aggregation Model, resources described in a imsmanifest.xml file, organized in schema/definition (.xsd and .dtd) files, and placed in a zip file are used as a content package. SCORM defines a web-based learning Content Aggregation Model and Run-Time Environment for learning objects. In SCORM, a content object is a web-deliverable learning unit. Often, a content object is just an HTML page or document that can be viewed with a web browser. A content object is the lowest level of granularity of learning resources, and can use all the same technologies a web page can use (e.g., Flash, JavaScript, frames, and images).

MPEG-21 [18] aims to define an open framework for multimedia applications, to support for example declaration (and identification), digital rights management, and adaptation. MPEG-21 is based on two essential concepts: the definition of a fundamental unit of distribution and transaction, which is the digital item, and the concept of users interacting with them. Within an item, an anchor binds descriptors to a fragment, which corresponds to a specific location or range within a resource. Items are grouped in a structured container using an XML-based Digital Item Declaration Language (DIDL). In addition a W3C XML Schema definition of DIDL is provided.

Table 2 summarizes OAI-ORE, METS, SCORM and MPEG-21 regarding basic principles available in complex objects: what is the data basic unit, how to relate a part of a document, how to identify it, and how to structure the components.

2.2 Superimposed-Information

In document creation, as in many other endeavors, re-use of information is often key to end-user productivity. Portions of a user's prior work, or that of other works, are often cited, inserted into, or otherwise used to enrich new works. Such activity is evident, for example, in the re-use of learning objects [64, 17], the use of annotations derived from prior works, and the preparation of teaching materials and derivative scholarly works. A number of authoring tools support such re-use by allowing a user to select a segment, or *subdocument* from a work and either enrich it (e.g., by annotation [11], by reference [**?**], or by using it elsewhere, for example using the copy-and-paste capability seen in most user interfaces). Unfortunately, most digital libraries (DLs) do not provide services supporting such use models. There is typically no facility for identifying or distinguishing subdocuments of interest from their enclosing documents. Further, there is no provision for a subdocument to have distinct metadata. As a result, subdocuments are not separately accessible, searchable, or manageable in most DLs. Thus information at a granularity important for frequent tasks is difficult to locate, understand, share, and use in many DLs. This motivates us to define and develop a **Superimposed-Information-Supported Digital Library** (SI-DL) with the goal of facilitating tasks that involve working with fine-grain contextualized information.

One foundation of our work is the notion of *Superimposed Information* (SI) [48, 23]: supplemental information created to reference, distinguish, extend, and organize subdocuments. SI existed long before digital information systems, but carries over into the digital world just as readily as other information forms. For example, citations and indexes are forms of *referential* SI: they allow users to reference or specify the location of subdocuments. The explicit highlighting of text, e.g., to label or tag it, is an example of *distinctional* SI: this allows subdocuments to be distinguished from surrounding material. Annotations and concordances as well as tags or labels are *extensional* SI: they augment and clarify the semantics of subdocuments. Finally, concept maps and multimedia presentations composed from existing information are examples of *organizational* SI: they organize collections of subdocuments into new works. Literature on digital forms of SI has included development and demonstration of infrastructure for the creation, resolution, and use (through development of superimposed applications) of SI [48, 51, 52, 53]. Superimposed applications may explicitly support any or all of these kinds of SI. As an example, an SI-enabled concept map tool [56] allows the user to associate subdocuments with a concept (using referential SI), organized into a concept map (using both organizational and extensional SI).

Central to digital SI is the notion of user (rather than author) identification of subdocuments of interest *in situ* (see Figure 2) with a *mark* [2]: an encapsulated address of a subdocument within its *base* (enclosing) document. The literature demonstrates mark capabilities both by storing marks directly in superimposed documents and by storing marks in a purpose-built repository within a middleware layer, facilitating mark browsing and re-use [51]. Also central to SI, in the context of a DL, is the notion of maintaining metadata for a subdocument separately from that of its base document, to enable searching and indexing of subdocuments. Various projects described in the literature have demonstrated these notions in the context of digital libraries.

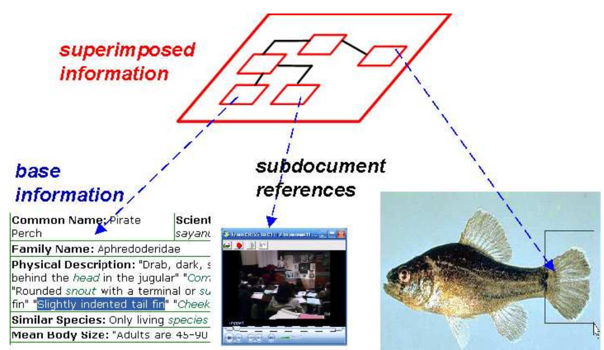

Fig. 2 *Working with information selections in situ.*

DL literature also contains substantive conceptual work on subdocument annotations [13, 7, 49, 81], a form of distinctional and extensional SI. Agosti [8, 6] proposes that annotations should be documents in their own right, so they can be browsed and searched independently. In the Digital Library for Earth System Education (DLESE) [11], annotations are stored as separate metadata records. However, annotations in DLESE are not explicit in the DL. Annotea

Name	Unit	Internal Component	Identifier	Structure
OAI-ORE	Resource	behaves like html	URI	Named Graph
METS	Simple object	FContent structure	OBJID	Structural Map
SCORM	Asset	sequence rules	——	schema/definition files
MPEG-21	Resource	anchors and fragments	URI	XML-DIDL

Table 2 How standards handle basic CO concepts.

[38] relies on markup clients that provide explicit support for annotation objects, and supports local and distributed repository storage of annotations. However, Annotea does not make sub-documents explicit in repositories, either.

Organizational SI is addressed in several places in the literature. Hypermedia models such as the Amsterdam model [34] extend the hypertext notion of links [33] to time-based media and compositions of different media. However, there is limited support in hypertext models and systems to work with subdocument information in situ, or with subdocuments defined by a user (rather than by the author). For example, in standards such as XLink and and XPath, sub-documents may typically be referenced, but only if pre-defined by the author, or if encompassed within XML tags [1, 72]. Superimposed documents and complex objects relate, also, to the idea of secondary repositories, where users may compose structured collections of complex digital objects [67]. These objects point back to the primary digital objects (similar to base information) from which they are produced. The focus of the project [67] is to examine the role of secondary repositories in access and preservation.

Organizational SI is also seen in work by Kerne et al., on recombinant information and hypersigns. This work focuses on developing compositions for visual semiotics[1] supporting personal expression to promote creative process and information discovery [39**?**]. The objective of knowledge management (KM) systems is to support creation, transfer, and application of knowledge in organizations [9]. SI offers a rich structuring opportunity that can be used for knowledge management. From the KM literature comes ideas of personal (or group) arrangements or organization of information to fulfill a task - a form of organizational SI.

Ted Nelson's Xanadu system presented two ideas – deep content links and transclusion, to describe his vision of hypertext (connected, networked documents), beyond what the World Wide Web implemented [63]. Transclusion, where primary information like quotations and annotations may be connected to subdocuments in their original context, is an example of distinctional and referential SI. In addition, transclusion can be viewed as a kind of extensional SI, in that the context surrounding the referenced subdocuments add meaning to the primary information.

More recently, Archer et al [10] defined and demonstrated an architecture for representing marks (i.e., subdocuments) as first-class items with metadata and annotation in the popular DSpace DL system [3], and have demonstrated the same capability in the Fedora DL system. We consider this work a form of distinctional and referential SI, though the inclusion of metadata constructs also touches on extensional SI, and the ability to organize marks into DL collections touches on organizational SI.

The focus of our work is on developing a formal representation for all forms of SI (distinctional, referential, extensional, and organizational) in a DL environment, and to study the use and reuse of such information in educational tasks such as teaching and learning.

2.3 CBIR

There are several digital libraries that support services based on image content [16, 82, 35, 27, 79, 80]. One example is the digital museum of butterflies [35], aimed at building a digital collection of Taiwanese butterflies. This digital library includes a module responsible for content-based image retrieval based on color, texture, and patterns. In a different image context, Zhu *et al.* [82] present a content-based image retrieval digital library that supports geographical image retrieval. The system manages air photos which can be retrieved through texture descriptors. Place names associated with retrieved images can be displayed by cross-referencing with a Geographical Name Information System (GNIS) gazetter. In this same domain, Bergman *et al.* describe an architecture for storage and retrieval of satellite images and video data from a collection of heterogeneous archives.

Another important initiative for digital library domain is related to the proposal of the Content-Based Image Search Component (CBISC) [76]. CBISC is a recently developed component that provides an easy-to-install search engine to query images by content. It can be readily tailored for a particular collection by a domain expert, who carries out a clearly defined set of pilot experiments. It supports the use of different types of vector-based image descriptors (metric and non-metric; color, texture, and shape descriptors; with different data structures to represent feature vectors), which can be chosen based on the pilot experiment, and then easily combined to yield improved effectiveness. The CBISC is an OAI-like search component which aims at supporting queries on image content. As in the OAI protocol [43],

[1] to construct and to understand new meanings

queries are submitted via HTTP requests. Two special requests ("verbs") are supported by this image search component: **ListDescriptors**, used to retrieve the list of image descriptors supported by our *CBISC*; and **GetImages**, used to retrieve a set of images by taking into account their contents.

Other initiatives cover different concepts explicated in the formalism presented below. For example, research presented in [27, 79] concentrates on new searching strategies for improving the effectiveness of CBIR systems, and another effort proposes image descriptors [80].

Several have worked to formalize content-based image retrieval systems [77, 12]. However, these formalisms typically describe these kinds of services under the database perspective (in general, based on the relational or object-relational models). To the best of our knowledge this paper constitutes the first formal attempt to describe content-based image retrieval services by using digital library concepts. One benefit is that the 5S framework is generic enough to formalize these services without relying on implementation decisions.

3 Complex Objects

From the computational view, a DL is composed of simple components named digital objects.

Recall the definition of a digital object [28]. A **digital object** is a tuple $do = (h, SM, ST, StructuredStreams)$, where

1. $h \in H$, where H is a set of universally unique handles (labels);
2. $SM = \{sm_1, sm_2, \ldots, sm_n\}$ is a set of streams;
3. $ST = \{st_1, st_2, \ldots, st_m\}$ is a set of structural metadata specifications;
4. $StructuredStreams = \{stsm_1, stsm_2, \ldots, stsm_p\}$ is a set of StructuredStream functions defined from the streams in the SM set (the second component) of the digital object and from the structures in the ST set (the third component).

COs are single entities that are composed of multiple digital objects, each of which is an entity in and of itself [41]. In other words, a complex digital object is a simple digital object or a recursive composition of other complex objects, as shown in Figure 3.

A complex digital object can be a digital object or an organization of other complex objects; therefore needing a structure to organize its components.

Definition 1 We define a complex digital object as a tuple $cdo = (h, SCDO = DO \cup SM, S)$ where

1. $h \in H$, where H is a set of universally unique handles (labels);

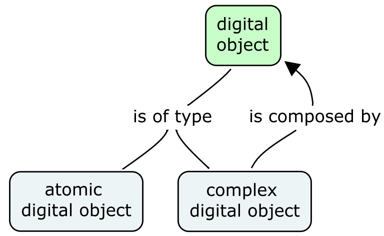

Fig. 3 A concept map for complex object composition.

2. $DO = \{do_1, do_2, \ldots, do_n\}$, where do_i is a digital object;
3. $SM = \{sm_1, sm_2, \ldots, sm_n\}$ is a set of streams;
4. S is a structure that composes the complex object cdo into its parts in $SCDO$.

A complex object is a simple digital object or a composition of other complex objects. The composition of its subparts (as seen in Figure 3) is represented by the component S.

This definition can also be used, for example, to represent a compound object cdo in OAI-ORE. The cdo could be represented as $cdo = (h, SCDO = DO \cup SM, S)$ where

1. $h \in H$, where H is a set of OAI-ORE URIs;
2. DO = $\{do_1, do_2, ..., do_n\}$, where do_i is a digital object;
3. SM = $\{sm_1, sm_2, ..., sm_n\}$ is a set of streams;
4. S is a structure that represents the same organization available in the OAI-ORE resource map.

An atomic digital object (mentioned in Figure 3) follows the same digital object definition as presented in [28].

4 A Superimposed-Information-Supported Digital Library

An SI-DL metamodel formally defines the various components that comprise an SI-DL. We extend the 5S minimal DL framework to include support for subdocuments, superimposed documents, and the relevant services. In terms of content, the main addition is the distinction among three types of digital objects: 1) base document – information existing as whole documents for which subdocuments have been defined; 2) subdocument – part of a base document referenced by an address into the base document; and 3) superimposed document – a separate document comprising of subdocuments and other information. It is important to highlight the temporal ordering that exists among the aforementioned types of digital objects, as depicted in Figure 4. The ordering relationship is similar to the temporal dimension of digital objects described by Agosti and Ferro in their formal model of annotations [6]. The temporal ordering states that a base document existed before a subdocument was created in it, which in turn, existed before or is created as it is used in a superimposed document. This limits the creation of a subdocument to the existence of its containing base document and limits the creation of a superimposed document to referencing existing subdocuments. Base documents, subdocuments, and superimposed documents, have all of the ordinary properties of a digital object as well, such as having metadata associated with it and being part of one or more collections. The content of each of these digital objects and their associated metadata can be browsed, indexed, and searched, as with any other digital object. In addition to existing services, we need a new service to deal with the referencing and presentation of a subdocument in situ. We call this service, *view in context*. The view in context service enables a subdocument to be viewed in the original context of its containing base document.

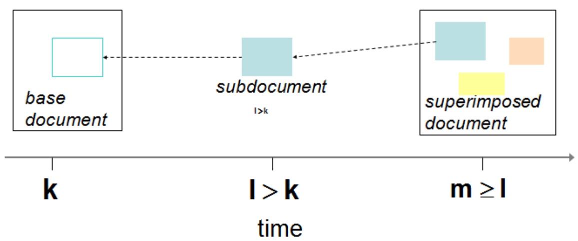

Fig. 4 *Temporal relationship among digital objects in an SI-DL.*

We assume that subdocuments and all kinds of superimposed information exist in the DL along with ordinary digital objects [2]. The activity of creation/composition is outside the scope of these definitions just as the authoring of digital objects is generally supported by tools that are outside of the DL. Thus, creating a subdocument, annotating a subdocument or another digital object, creating/composing a superimposed document, such as a concept map, strand map, etc. are all outside of the scope of our model. We are only concerned with how this information is represented in a DL and what new services will be added to access, retrieve, and facilitate viewing of information once is has been added to the DL. Note that specific superimposed applications are responsible for viewing superimposed documents and the SI-Dl formalization is not concerned with those applications[3].

We need to make a comment about *annotation* here. It is an important part of an SI-DL since it is supplemental information associated with a subdocument. However, an annotation may be associated with any kind of digital object as well and is not restricted to subdocuments. We choose to describe an annotation as a new superimposed document consisting of the text or other material comprising the annotation (or link to a digital object comprising the annotation) that references a subdocument or other document, i.e., the original material in a base document) that is being annotated. Note that every superimposed document includes the address of any subdocuments that it references. This is in line with the formal definition of annotation by Agosti and Ferro [6]. Thus, we do not define annotation explicitly in our metamodel.

The new concepts added to a DL are as shown in the Figure 5. The figure also shows the connection between a superimposed document and a complex object. In the remaining part of this section, we formally define the components of an SI-DL.

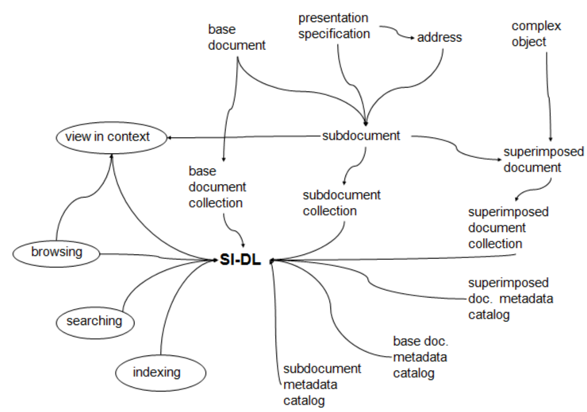

Fig. 5 *Definitional dependencies among concepts in an SI-DL.*

4.1 5S Extensions

4.1.1 Base Document

A **base document** *BD* is a digital object for which a subdocument exists. Any digital object can thus become a *BD*, upon creation of the first subdocument. See Section 3, which has a review of the definition of a digital object[4].

4.1.2 Presentation Specification, Address, and Subdocument

In this section, we define all concepts associated with a subdocument. We extend the definitions of substream in the 5S framework [31] and segment in the formal annotation model [6] to define a subdocument. According to Goncalves et al., a segment or substream is associated with a pair of natural numbers $(a,b), a < b$, corresponding to a contiguous subsequence $[S_a, S_b]$ of stream S. Or, we can say $sm_t[i,j] = \langle a_i, a_1, \ldots, a_j \rangle$, $0 \le i \le j \le n$ is a substream or segment of stream S. According to Agosti and Ferro, given a stream sm: $I = \{1,2,...,n\} \rightarrow \Sigma$, where Σ is the alphabet of symbols and $n \in N, sm \in SM$, a segment is a pair: $st_{sm} = (a,b)$ such that $1 \le a \le b \le n, ab \in N$.

[2] Ordinary digital objects need not be either of: a base document, a subdocument, or a superimposed document.

[3] In a similar way that we are not concerned about display of base documents.

[4] for details, see definition 16 in the 5S framework [31]

In addition to getting the content of the base document that comprises the subdocument, we need to retain the base document context of the subdocument (to allow tools to view /present it *in situ*). We do so by extending the aforementioned definitions of substream and segment to include *presentation specification* and *address*. Also, we store other associated information with a subdocument including properties (such as its creator and timestamp of its creation) and semantic attributes (such as annotations and tags) as part of promoting the subdocument to be a first-class concept within a digital library.

Presentation specification provides information about how a subdocument was defined in a base document. This notion is borrowed from the hypertext/hypermedia world, where it refers to the runtime behavior of information units presented to the user [33, 34]. In hypertext/hypermedia literature, presentation specification refers to the encoding information and mechanism that is used to present a component (or network of components) to the user. A software application/tool uses the presentation specification to display the contents of a digital object. A presentation specification is a descriptive metadata specification conforming to a presentation-based metadata format (definition 13 and 14 in [31]). It is used to specify how the content in a digital object translates into a particular view/presentation. Presentation specification includes information such as the content type of the base document (text, image, audio, video, etc.), the format of the base document (.PDF, .DOC, .JPEG, .AVI, etc.), and the specific software tool used to view/present the base document(Adobe Acrobat, Microsoft Word, Microsoft image viewer, etc.), used when the subdocument was created.

Definition 2 A **presentation specification**, $PS = (G_{PS}, \mathscr{R}_{PS} \cup \mathscr{L}_{PS} \cup \mathscr{P}_{PS}, \mathscr{F}_{PS})$ **conforms with** a presentation-based metadata format $MF_{PS} = (V_{MF_{PS}}, \mathrm{def}_{MF_{PS}})$ with the following constraints:

1. $V_{MF_{PS}} = \{\mathscr{R}_{PS1}, \mathscr{R}_{PS2}, ..., \mathscr{R}_{PSk}\} \subset 2^{\mathscr{R}_{PS}}_{MF}$ a family of subsets of the resources labels $\mathscr{R}_{PSMF}$ and $\mathrm{def}_{MF_{PS}} : V_{MF_{PS}} \times \mathscr{P}_{MF_{PS}} \rightarrow V_{MF_{PS}} \cup D_{\mathscr{L}_{MF_{PS}}}$ is a property definition function.
2. $\mathscr{R}_{PS} \subseteq \mathscr{R}_{MF_{PS}}$,
3. $\mathscr{L}_{PS} \subseteq \mathscr{L}_{MF_{PS}}$,
4. $\mathscr{P}_{PS} \subseteq \mathscr{P}_{MF_{PS}}$, and
5. for every statement $st = (r, p, l)$ derived from PS, $r \in \mathscr{R}_k$ for some $\mathscr{R}_k \in V_{MF_{PS}}$ and $p \in \mathscr{P}_{PS}$ implies $l \in \mathrm{def}_{MF_{PS}}(\mathscr{R}_k, p)$.
6.

Example for resources could be an academic paper, an image, a software application, etc. Examples of properties include format, content type, software application to view, etc. Consider the example shown in Figure 6. Here the object "Shield Darter" is an "image" of "JPEG" format and makes use of the "Java image viewer" software application. Another example is from the Dublin Core metadata format. For any set of labels $\mathscr{R}$ for resources, the Dublin Core metadata format defines that $\mathrm{def}_{DC}(\mathscr{R}, `format') = String$ and $\mathrm{def}_{DC}(\mathscr{R}, `format.mimetype') = MIME$ where $MIME$ is a finite set of labels for Resources corresponding to mime types.

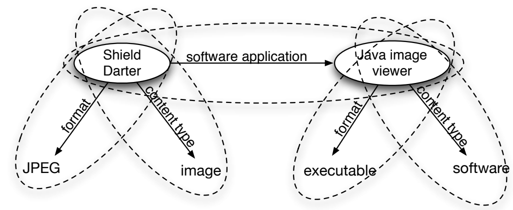

Fig. 6 *Example of a presentation specification.*

A presentation specification of a subdocument is used to get the *address* of the span/region of the subdocument within the base document. The address is, then, used by an appropriate software application to navigate to and view the subdocument in context of its originating base document. Consider the example of an academic paper, which might have mixed content including text and images. It could be a PDF document presented/viewed using Adobe Acrobat. The address of a segment or substream in this case might be different than if the same content were in a .DOC document presented/viewed using Microsoft Word since the navigation/addressing schemes within each of these tools is different. Adobe Acrobat uses a word-based scheme whereas Microsoft Word uses a character-based scheme. Another example is the address of a subdocument within an image document (or a subimage), which might vary depending on the format, resolution, and software used to view/present the image. In our current work that extends upon previous work [**?**] on including subdocuments in the DSpace DL software [3], we have implemented a feature for Microsoft Word (and OpenOffice) that allows for creation of subdocuments (which we have stored in an instance of the Fedora DL [4]) and also are able to accept an address for a subdocument with a Microsoft Word (and OpenOffice) base document and display it highlighted.

Definition 3 Given base document *BD*, a **subdocument** *sd* is a digital object with the following extensions and constraints:

– *sd* is a *digital object* = (h, SM, ST, $StrStreams$, PS, $addr$), where
  1. $h \in H$, where H is a set of universally unique handles (labels);

2. $SM_{sd} = \{sm_{sd}[i,j]\} \in SM$, where $sm_{sd}[i,j] = \langle a_i, \ldots, a_j \rangle, 0 \leq i \leq j \leq n$. $sm_{sd}[i,j]$ refers to substreams of a base document *BD*.
3. $ST = \{st_1, st_2, \ldots, st_m\}$ is a set of structural metadata specifications associated with the base document *BD*;
4. $StrStreams = \{stD_1, stD_2, \ldots, stD_m\}$ is a set of StructuredStream functions defined from the base document substreams in the SM_{sd} set (the second component) of the subdocument and from the structures in the ST set (the third component).
5. *PS* is a *presentation specification*.
6. *addr* is the function from the SM_{sd} set (the second component) of the subdocument and from the presentation specification *PS* of the base document.

Note that the subdocument contains the *structures* and the contiguous *streams* and of its parent base document that exist within the span defined by the address of the subdocument. It inherits all the descriptive and structural metadata specifications associated with the span defined by the address. Figure 7 shows an example of a subdocument with its components, including the substreams and substructures associated with it, as inherited from the containing base document. In addition, it has an address that is a function of the presentation specification, *PS* associated with the subdocument. Since a subdocument is a digital object, it has its own metadata. This could include properties of subdocument creation such as information about the subdocument creator, the timestamp of creation, etc. Also, as with an ordinary digital object, a subdocument could be associated with semantic information such as annotations and tags. Like other digital objects, a subdocument may have many manifestations. For example, consider a subdocument within a text-based PDF document. One manifestation of the subdocument might be the textual excerpt of the subdocument. Another might be an image transformation of a portion of the base PDF document with the highlighted subdocument.

4.1.3 Superimposed Document

A superimposed document can be represented as a complex object (as defined in section 3), where at least one of its constituent digital objects is a subdocument.

Definition 4 A superimposed document is a complex digital object, defined as a tuple
$sidoc = (h, DO \cup SM, S, ST)$, where

1. $h \in H$, where H is a set of universally unique handles (labels);
2. DO = $\{do_1, do_2, ..., do_n\}$ is a set of digital objects that are part of the superimposed document, *sidoc*, such that $\exists$ at least one $do_i = sd$, for $i = 1, 2, \cdots, n$, were sd is a subdocument;
3. $SM = \{sm_1, sm_2, \ldots, sm_n\}$ is a set of streams;
4. S is a structure that composes the superimposed document *sidoc* from its component parts in $DO \cup SM$.

This is consistent with our earlier work in SI where the references to subdocuments (i.e., marks) could be incorporated into a variety of superimposed documents structured according to various data models [54]. A superimposed document can be of different types. For example, it may consist of subdocument references (i.e., marks) interspersed with other digital content, such as in a textual document that has citations to specific portions of other documents. Another example is a time-ordered arrangement of audio/video clips merged with textual content from web pages [55]. A concept map [56] or a strandmap [22], where the resources point to subdocuments are other examples.

4.2 Collections and Catalogs

A key component of a digital library with SI support is the ability to deal with collections and metadata catalogs. Here, we define collections and catalogs for the three types of digital objects that we have introduced.

Definition 5 A **base document collection**
$C_{BD} = \{bd_1, bd_2, \ldots, bd_k\}$ is a set of base documents.

Definition 6 A **subdocument collection**
$C_{sd} = \{sd_1, sd_2, \ldots, sd_k\}$ is a set of subdocuments.

Definition 7 A **superimposed document collection**
$C_{sidoc} = \{sidoc_1, sidoc_2, \ldots, sidoc_k\}$ is a set of superimposed documents.

Definition 8 Let C_{BD} be a base document collection with k handles in H. A **base document metadata catalog** $DM_{C_{BD}}$ for C_{BD} is a set of pairs
$\{(h, \{dm_{BD_1}, \ldots, dm_{BD_{k_h}}\})\}$, where $h \in H$ and the dm_{BD_i} are descriptive metadata specifications for *BD*, the base document.

Definition 9 Let C_{BD} be a subdocument collection with k handles in H_{sd}. A **subdocument metadata catalog** $DM_{C_{sd}}$ for C_{sd} is a set of pairs
$\{(h, \{dm_{sd_1}, \ldots, dm_{sd_{k_{h_{sd}}}}\})\}$, where $h_{sd} \in H_{sd}$ and the dm_{sd_i} are descriptive metadata specifications for the subdocument, *sd* .

Definition 10 Let C_{BD} be a superimposed document collection with k handles in H. A **superimposed document metadata catalog** $DM_{C_{sidoc}}$ for C_{sidoc} is a set of pairs
$\{(h, \{dm_{sidoc_1}, \ldots, dm_{sidoc_{k_h}}\})\}$, where $h \in H$ and the dm_{sidoc_i} are descriptive metadata specifications for superimposed document, *sidoc*.

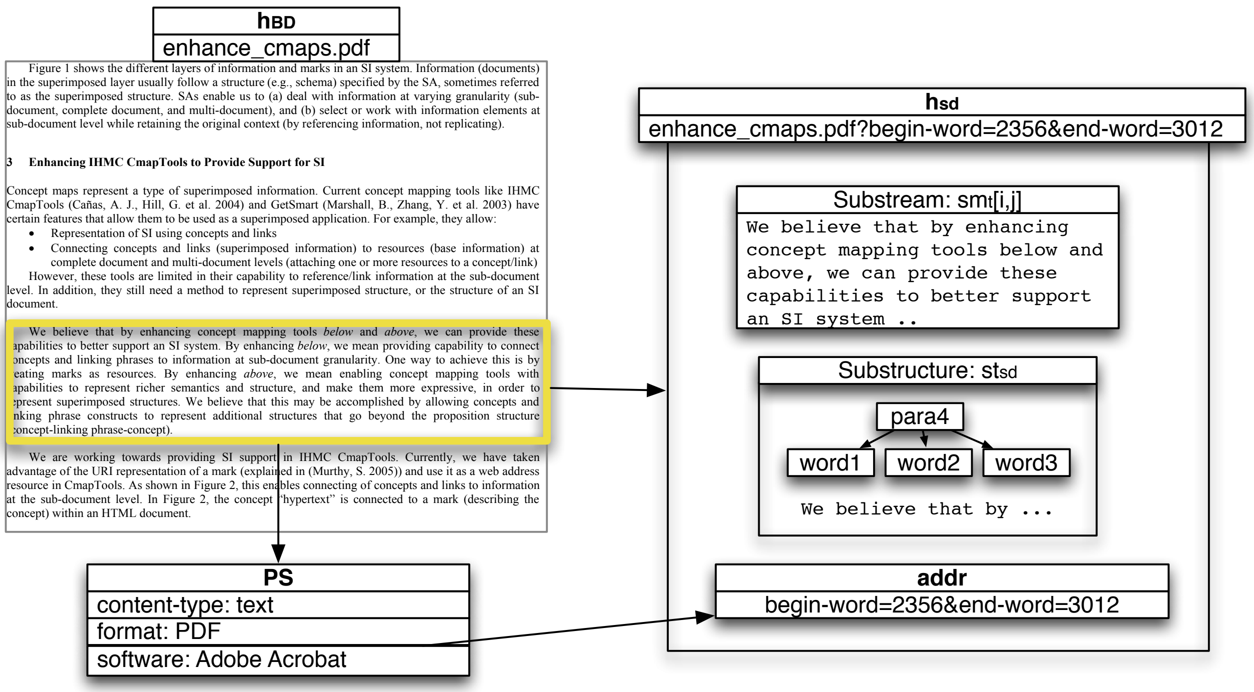

Fig. 7 *Example of a subdocument and its components.*

4.3 Services

Traditional services such as browsing, indexing, and searching will now act on different types of digital objects including base documents, subdocuments, superimposed documents, as well as metadata associated with each of these, including marks. For example, using the search service on subdocuments, the query specification can contain subdocument -related information and the results can include subdocuments. In addition, advanced searches on components of superimposed documents and base documents might be possible. For example, one could get all subdocuments within a particular base document. Another example is all base documents that contain subdocuments, which are referenced in a superimposed document.

In addition to traditional digital library services, a new service, *view in context* is added to the digital library to support access for viewing/presentation of subdocuments in the context of their parent base document. This can be considered an extension of the browsing services as defined in the 5S framework, which acts upon the extended hypertext that now includes subdocuments and links between base documents and subdocuments as well as those between superimposed documents and subdocuments. This creates new referential hyperlinks between a subdocument and its parent document as well as those between a superimposed document and its constituent subdocuments. In addition, we now need to make use of links to services, for example plugins that can be invoked by the digital library based on the presentation specification of the base document which contains a subdocument.

Definition 11 A *view in context service* is a set of scenarios $\{sc_1,\ldots,sc_n\}$ over a an extended hypertext where events are defined by edges of the hypertext graph (V_{H_E}, E_{H_E}), where V_{H_E} includes the union of base documents, subdocuments, and superimposed documents and E_{H_E} includes the links between a subdocument and base document, such that the subdocument–base document link events e_i are associated with a function $ViewInContext : V_{H_E} \times E_{H_E} \rightarrow Contents$, which given a subdocument, instantiates the service that is required to present/view the base document (facilitated through information in the presentation specification, *PS*), retrieves the content of the base document and uses the aforementioned service for the base document's presentation with the subdocument highlighted within the base document, i.e., $ViewInContext(v_{k_{sd}}, e_{k_i}) = \mathscr{P}(v_{t_{sd}})$ for $e_{k_i} = (v_{k_{sd}}, v_{t_{sd}}) \in E_{H_E}$. Here, $v_{k_{sd}}$ is a reference of the subdocument in the superimposed layer of information and $v_{t_{sd}}$ is the subdocument in it original context in the base layer of information.

An example of the view in context service is shown in Figure 8. Here, the subdocument used in the superimposed layer is created from a Microsoft Word document with a plugin that allows subdocument creation and viewing. On instantiating the view in context service from this subdocument, an instance of Microsoft Word is launched, the base

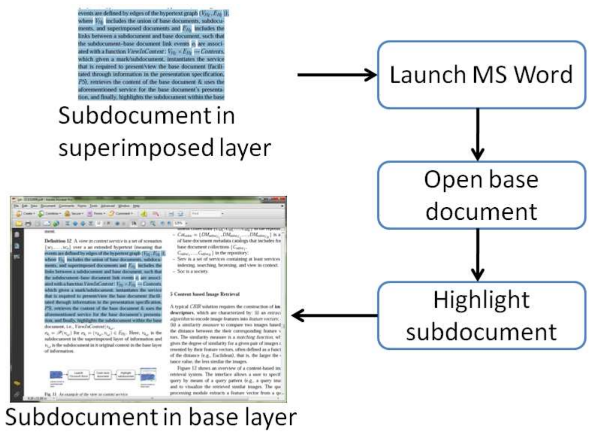

Fig. 8 *An example of the view in context service.*

document containing the subdocument is opened and presented in the Word application, and the subdocument is highlighted in this base document.

4.4 SI-DL

Definition 12 A **superimposed information supported digital library** is a 4-tuple $(\mathscr{R}, DM, Serv, Soc)$, where

- $\mathscr{R}$ is a repository;
- $DM = DM_{BD} \cup DM_{sd} \cup DM_{sidoc} \cup DM_{do}$,
- $DM_{BD} = \{DM_{BD_{C_1}}, DM_{BD_{C_2}}, \ldots, DM_{BD_{C_K}}\}$ is a set of base document metadata catalogs for all base document collections $\{C_{BD_1}, C_{BD_2}, \ldots, C_{BD_K}\}$ in the repository;
- $DM_{sd} = \{DM_{sd_{C_1}}, DM_{sd_{C_2}}, \ldots, DM_{sd_{C_K}}\}$ is a set of subdocument metadata catalogs for all subdocument collections $\{C_{sd_1}, C_{sd_2}, \ldots, C_{sd_K}\}$ in the repository;
- $DM_{sidoc} = \{DM_{sidoc_{C_1}}, DM_{sidoc_{C_2}}, \ldots, DM_{sidoc_{C_K}}\}$ is a set of base document metadata catalogs for all base document collections $\{C_{sidoc_1}, C_{sidoc_2}, \ldots, C_{sidoc_K}\}$ in the repository;
- DM_{do} is a set of metadata catalogs for all collections $\{C_{do_1}, C_{do_2}, \ldots, C_{do_K}\}$ in the repository, that are not in the sets of base document, subdocument, and superimposed document collections;
- Serv is a set of services containing at least services for indexing, searching, browsing, and view in context;
- Soc is a society.

5 Content-based Image Retrieval

A typical *CBIR* solution requires the construction of **image descriptors**, which are characterized by: (i) an *extraction algorithm* to encode image features into *feature vectors*; and (ii) a *similarity measure* to compare two images based on the distance between the their corresponding feature vectors. The similarity measure is a *matching function*, which gives the degree of similarity for a given pair of images represented by their feature vectors, often defined as a function of the distance (e.g., Euclidean), that is, the larger the distance value, the less similar the images.

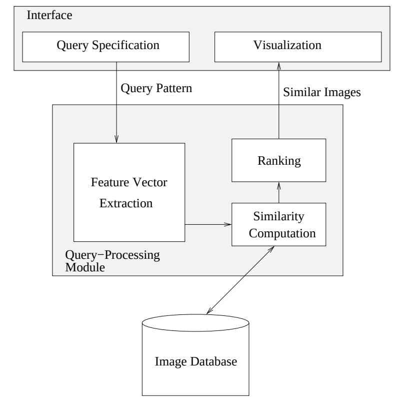

Fig. 9 Typical CBIR system.

Figure 9 shows an overview of a content-based image retrieval system. The interface allows a user to specify a query by means of a query pattern (e.g., a query image) and to visualize the retrieved similar images. The query-processing module extracts a feature vector from a query pattern and applies a distance function (such as the Euclidean distance) to evaluate the similarity between the query image and the images. Next, it ranks the database images according to similarity and forwards the most similar images to the interface module. Note that database images are often indexed according to their feature vectors using structures to speed up retrieval and distance computation.

5.1 5S Extensions

Figure 10 presents the proposed concepts based on the 5S framework to handle image content descriptions and related digital library services. These concepts are precisely defined below.

Some of these concepts were introduced in [74]. In this paper, we extend them by taking into account digital library aspects.

Definition 13 An **image *stream*** (or simply **image**) $\hat{I}$ is a pair $(D_I, \mathbf{I})$, where:

- D_I is a finite set of *pixels* (points in $\mathbb{N}^2$, that is, $D_I \subset \mathbb{N}^2$), and
- $\mathbf{I} : D_I \rightarrow \mathrm{D}'$ is a function that assigns each pixel p in D_I to a vector $\mathbf{I}(p)$ of values in some arbitrary space D' (for example, $\mathrm{D}' = \mathrm{I\!R}^3$ when a color in the RGB system is assigned to a pixel).

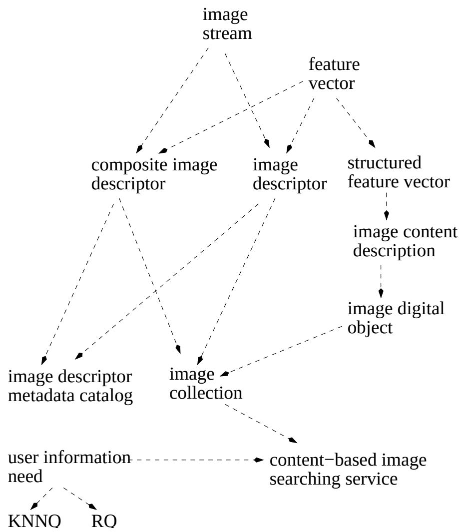

Fig. 10 5S extensions to support content-based image description and related services.

Definition 14 A **feature vector** $\mathbf{fv}_{\hat{I}}$ of an image $\hat{I}$ is a point in $\mathbb{R}^n$ space: $\mathbf{fv}_{\hat{I}} = (fv_1, fv_2, ..., fv_n)$, where n is the dimension of the vector.

Examples of possible feature vectors are a color histogram [71], a multiscale fractal curve [75], and a set of Fourier coefficients [65]. They essentially encode image properties, such as color, shape, and texture. Note that different types of feature vectors may require different similarity functions.

Definition 15 Given a structure $(G, L, \mathscr{F})$, $G = (V, E)$ and a feature vector $\mathbf{fv}_{\hat{I}}$, a **StructuredFeatureVector** is a function $V \to \mathbb{R}^n$ that associates each node $v_k \in V$ with $fv_i \in \mathbf{fv}_{\hat{I}}$.

Figure 11 presents an example of the use of a **StructuredFeatureVector** function. In this case, an XML structure (structural metadata specification) is mapped to a feature vector obtained by applying the image descriptor *Contour Multiscale Fractal Dimension* [75] to the image stream defined by the file "fish0.pgm".

Definition 16 A **simple image content descriptor** (briefly, **image descriptor**) D is defined as a tuple $(h_{desc}, \varepsilon_D, \delta_D)$, where:

- $h_{desc} \in H$, where H is a set of universally unique handles (labels);
- $\varepsilon_D : \{\hat{I}\} \to \mathbb{R}^n$ is a function, which extracts a *feature vector* $\mathbf{fv}_{\hat{I}}$ from an *image* $\hat{I}$.
- $\delta_D : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a *similarity function* (e.g., based on a distance metric) that computes the similarity between two images as a function of the distance between their corresponding *feature vectors*.

Figure 12(b) illustrates the use of a simple descriptor D to compute the similarity between two images $\hat{I}_A$ and $\hat{I}_B$.


```
<?xml version="1.0" encoding="UTF−8"?>
−<feature_vector:Feature_Vector xmlns:feature_vector="http://feathers.dlib.vt.edu/~rtorres/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema−instance"
    xsi:schemaLocation="http://feathers.dlib.vt.edu/~rtorres/
    http://feathers.dlib.vt.edu/~rtorres/feature_vector.xsd">
    <feature_vector::ImageName>fish0.pgm</feature_vector:ImageName>
    <feature_vector:DescriptorName> ContourMSFractalDimension <feature_vector:DescriptorName>
    <feature_vector:Type> 1 <feature_vector:Type>
    −<feature_vector:Curve>
        <feature_vector:Nelements> 25 <feature_vector:Nelements>
        −<feature_vector:Curve1D>
            −<feature_vector:X>
                <feature_vector:value> 0.95105259594482394192 <feature_vector:value>
                <feature_vector:value> 0.98551214588154611995 <feature_vector:value>
                <feature_vector:value> 1.00415492765507829986 <feature_vector:value>
                <feature_vector:value> 1.00931032237937512441 <feature_vector:value>
                <feature_vector:value> 1.00583781572741104426 <feature_vector:value>
                                    (...)
                <feature_vector:value> 0.93810555611087775851 <feature_vector:value>
                <feature_vector:value> 0.87275204902189629230 <feature_vector:value>
                <feature_vector:value> 0.81066432563100665476 <feature_vector:value>
                <feature_vector:value> 0.75224263059381879515 <feature_vector:value>
            </feature_vector:X>
        </feature_vector:Curve1D>
    </feature_vector:Curve>
</feature  vector:Feature  Vector>
```


Fig. 11 Example of a structured feature vector.

First, the extraction algorithm ε_D is used to compute the feature vectors $\mathbf{fv}_{\hat{I}_A}$ and $\mathbf{fv}_{\hat{I}_B}$ associated with the images. Next, the similarity function δ_D is used to determine the similarity value d between the images.

Definition 17 A **composite image descriptor** $\hat{D}$ is a tuple $(h_{desc}, \mathscr{D}, \delta_{\mathscr{D}})$ (see Figure 12(b)), where:

- $h_{desc} \in H$, where H is a set of universally unique handles (labels);
- $\mathscr{D} = \{D_1, D_2, \ldots, D_k\}$ is a set of k pre-defined simple image descriptors.
- $\delta_{\mathscr{D}}$ is a similarity function which combines the similarity values obtained from each descriptor $D_i \in \mathscr{D}$, $i = 1, 2, \ldots, k$.

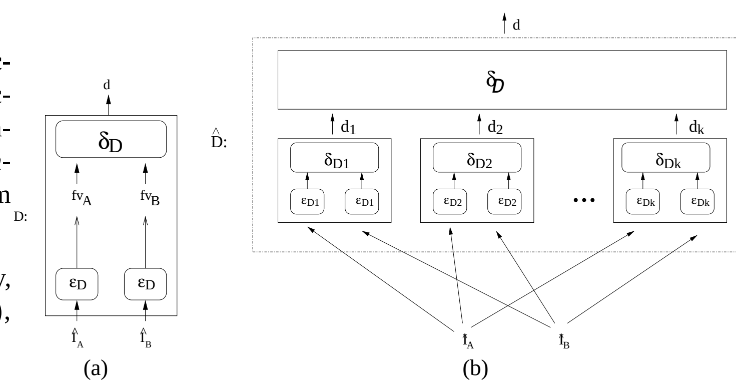

Fig. 12 (a) The use of a simple descriptor D for computing the similarity between images. (b) Composite image descriptor.

Definition 18 An **image content description** ICD is a tuple $(FV, ST_{FVs}, Structured_{FVs})$, where

- $FV = \{\mathbf{fv_1}, \mathbf{fv_2}, \ldots, \mathbf{fv_k}\}$ is a set of feature vectors;
- $ST_{FVs} = \{stfv_1, stfv_2, \ldots, stfv_m\}$ is a set of structural metadata specifications;

– $Structured_{FVs} = \{strfv_1, strfv_2, \ldots, strfv_m\}$ is a set of StructuredFeatureVector functions defined from the *feature vectors* in the *FV* set (the first component) of the image content description and from the structures in the ST_{FVs} set (the second component).

Definition 19 An **image digital object** *ido* is a digital object with the following extensions and constraints:

– *ido* is a *digital object* = (*h*, *SM*, *ST*, *StrStreams*, *ICD*, *StrICDStreams*), where
  1. $h \in H$, where H is a set of universally unique handles (labels);
  2. $SM_{sd} = \{sm_{sd}[i,j]\} \in SM$, where $sm_{sd}[i,j] = \langle a_i, \ldots, a_j \rangle, 0 \leq i \leq j \leq n$. $sm_{sd}[i,j]$ refers to substreams (regions) of an image stream.
  3. $ST = \{st_1, st_2, \ldots, st_m\}$ is a set of structural metadata specifications;
  4. $StrStreams = \{stD_1, stD_2, \ldots, stD_m\}$ is a set of StructuredStream functions defined from the image substreams in the *SM* set (the second component) of the digital object and from the structures in the *ST* set (the third component).
  5. *ICD* is an *image content description*.
  6. $StrICDStreams = \{stimgD_1, stimgD_2, \ldots, stimgD_m\}$ is a set of StructuredStream functions defined from the *image stream* in the *SM* set (the second component) of the image digital object and from the structures in the $ST_{FVs} \in ICD(2)$ set.

Figure 13 illustrates the relations among the concepts used to define an image digital object.

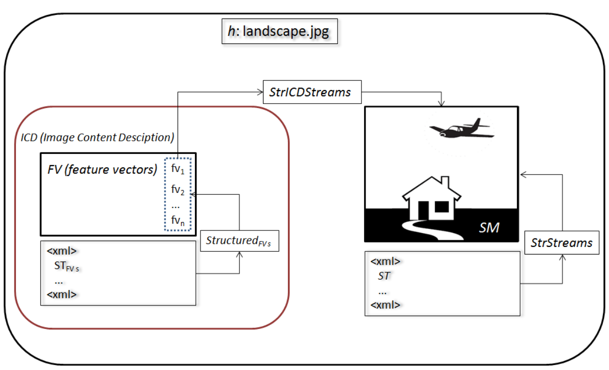

Fig. 13 Image digital object elements.

The definition of *StrICDStreams* allows associating feature vectors to parts (objects, regions) of image streams.

Definition 20 An **image collection** *ImgC* is a tuple $(C, S_{imgdesc}, FV_{imgdesc})$, where *C* is a collection (see Def. 17 in [31]), $S_{imgdesc}$ is a set of image descriptors, and FV_{desc} is a function $FV_{desc} : \{C \times S_{imgdesc}\} \rightarrow ICD(1)$, where *ICD* is $ido(5)$ and $ido \in C$.

Function FV_{desc} defines how a feature vector was obtained, given an image digital object $ido \in C$ and an image descriptor $\hat{D} \in S_{imgdesc}$.

Definition 21 Let $S_{imgdesc}$ be a set of image descriptors with *k* handles in *H*. An **image descriptor metadata catalog** $DM_{S_{imgdesc}}$ for $S_{imgdesc}$ is a set of pairs $\{(h, \{dmdesc_1, \ldots, dmdesc_{k_h}\})\}$, where $h \in H$ and the $dmdesc_i$ are descriptive metadata specifications for image descriptors.

Descriptive metadata specifications of descriptors could include, for example, data about the author (who implemented the extraction and similarity functions), implementation date, and related publication.

Recall that, in general, a metadata catalog is used to assign descriptive metadata specifications to image digital objects (see Def. 18 in [31]).

Definition 22 A conceptual representation for user information need is materialized into a query specification. A **query specification** *Q* is a tuple $Q = \{(H_q, Contents_q, P_q)\}$, where $H_q = ((V_q, E_q), L_q, \mathscr{F}_q)$ is a structure (i.e., a directed graph with vertices V_q and edges E_q, along with labels L_q and labeling function $\mathscr{F}_q$ on the graph; see Def. 2 in [31] for details), $Contents_q$ includes digital objects and all of their streams, and P_q is a mapping function $P_q : V_q \rightarrow Contents_q$.

The notion of conceptual representations for user information needs was used in [31] to define a searching service, however, it was not formally defined. The formal definition for conceptual representations for user information needs was originally presented in [68].

Usually, two kinds of queries are supported by CBIR systems [21]. In a *K-nearest neighbor query (KNNQ)*, the user specifies the number *k* of images to be retrieved that are closest to the query pattern. In a *range query (RQ)*, the user defines a search radius *r* and wants to retrieve all database images whose distance to the query pattern is less than *r*. In this case, both the specification of *k* in the KNNQ and the specification of *r* needs to be incorporated into *Q*.

Definition 23 A query specification $q \in Q$ is a **K-nearest neighbor query (KNNQ) information need** if there exists $v \in V_q$, a real number $k \in Contents_q$, and $P_q(v) = k$.

Definition 24 A query specification $q \in Q$ is a **range query (RQ) information need** if there exists $v \in V_q$, a real number $r \in Contents_q$, and $P_q(v) = r$.

Definition 25 A **content-based image searching service** is a set of searching scenarios $\{sc_1, sc_2, \ldots, sc_t\}$, where each scenario sc_i is a sequence of events, and each event e_i is associated with the OP_s function defined as follows:

$OP_s : (Q \times C) \times Sim_s \rightarrow 2^{Contents}$, where $Sim_s = OP_q(q, ido) | q \in Q, ido \in C$, and where $OP_q : Q \times C \rightarrow \mathbb{R}$ is a matching function that associates a real number with $q \in Q$ and a digital object $ido \in C$. The computation of OP_q relies on the use of appropriate image descriptors (e.g., their extraction and distance computation algorithms) defined in the image collection *ImgC*.

The range of function OP_s is the *Contents* associated with collection *ImgC*. While the similarity function OP_q was defined in Def. 21 in [31], the retrieved results were not defined there. We consider the retrieved results as (a subset of) the *Contents*.

6 Case Study

In this case study, we use the 5S extensions to define and analyze content and behavior of an image description and retrieval tool. The components of this tool are not unlike those of a digital library with extended functionality.

6.1 Superimposed Image Description and Retrieval Tool - SuperIDR

SuperIDR is a superimposed image description and retrieval tool [57, 60, 59], developed with the aim of helping users to work with parts of images in situ, where they can select, annotate, and retrieve parts of images in the context of the original image. We use *fish species identification* as the specific scholarly task to test the use of this tool. However, the tool might be used in any task involving images with a significant number of important details, such as analyzing paintings in art history, examining a building style in architecture, understanding trees in dendrology, etc.

SuperIDR is seeded with details of 207 species of freshwater fishes of Virginia, taken from [36]. Each species has a representative image as shown in Figure 14-b. In addition to making annotations, SuperIDR allows searching and browsing of species descriptions, images, image marks, and annotations. A user can search in one of two ways: 1) perform text-based search (full-text and field-wise search, powered by Lucene.NET[5]) on species descriptions and annotations, where the query may include boolean combinations of terms, phrases; 2) perform content-based image search on images and annotated-image-marks, where the query may be a complete image or part of an image. Finally, in SuperIDR, a user can browse through species information either through a taxonomic organization of species based on family and genera or through an electronic version of the dichotomous key. Scenario 5 describes the use of this tool by an Ichthyology student. Figure 14 shows screenshots of the tool.

Figure 15 shows the architecture of the SuperIDR tool. In the data layer, there are images, associated feature vectors, image subdocuments, associated annotations, fish species descriptions, data related to taxonomic classification of species , and identification key data. In the processing layer, we model the the annotation, search, and browse functionalities. The content-based image search is enabled by CBISC, a content-based image search component [73] and the text search is enabled by the Lucene search engine. The presentation layer contains interfaces for annotation, three types of searching, and browsing.

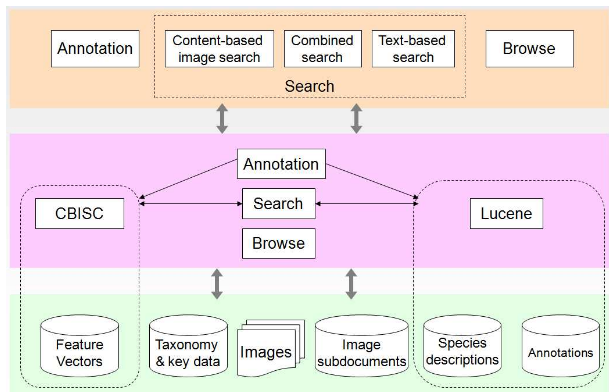

Fig. 15 Architecture of the of the SuperIDR tool.

6.2 Formalizing the Content and Functionality of SuperIDR

The SuperIDR digital library (SuperIDR DL) might be considered to be an extension of the minimal digital library as defined in the 5S framework [31]. Figure 16 shows the components of the SuperIDR DL. We have extended the definition of a digital object to include an image digital object, an image subdocument, a species complex object, a species superimposed complex object. In addition, SuperIDR has other digital objects, such as annotation and species description. These conform to the digital object definition as mentioned in the 5S framework. Each of the aforementioned digital object belongs to respective collections and is associated with a metadata catalog. In addition, SuperIDR has the view in context and CBIR services. The rest of this section formally describes the components of SuperIDR.

Figure 17 shows the information components within SuperIDR and relationships among them. Here, species is considered to be a complex object and it consists of at least one or more image digital objects and species description. When at least one of the images gets marked for annotation, a subdocument is created and added to the species digital object. Also, the associated annotation object is added to species. The addition of a subdocument makes species a superimposed complex object. Each of the aforementioned digital objects, image, image subdocument, annotation, species description, and species, has an associated metadata record.

[5] http://incubator.apache.org/lucene.net/

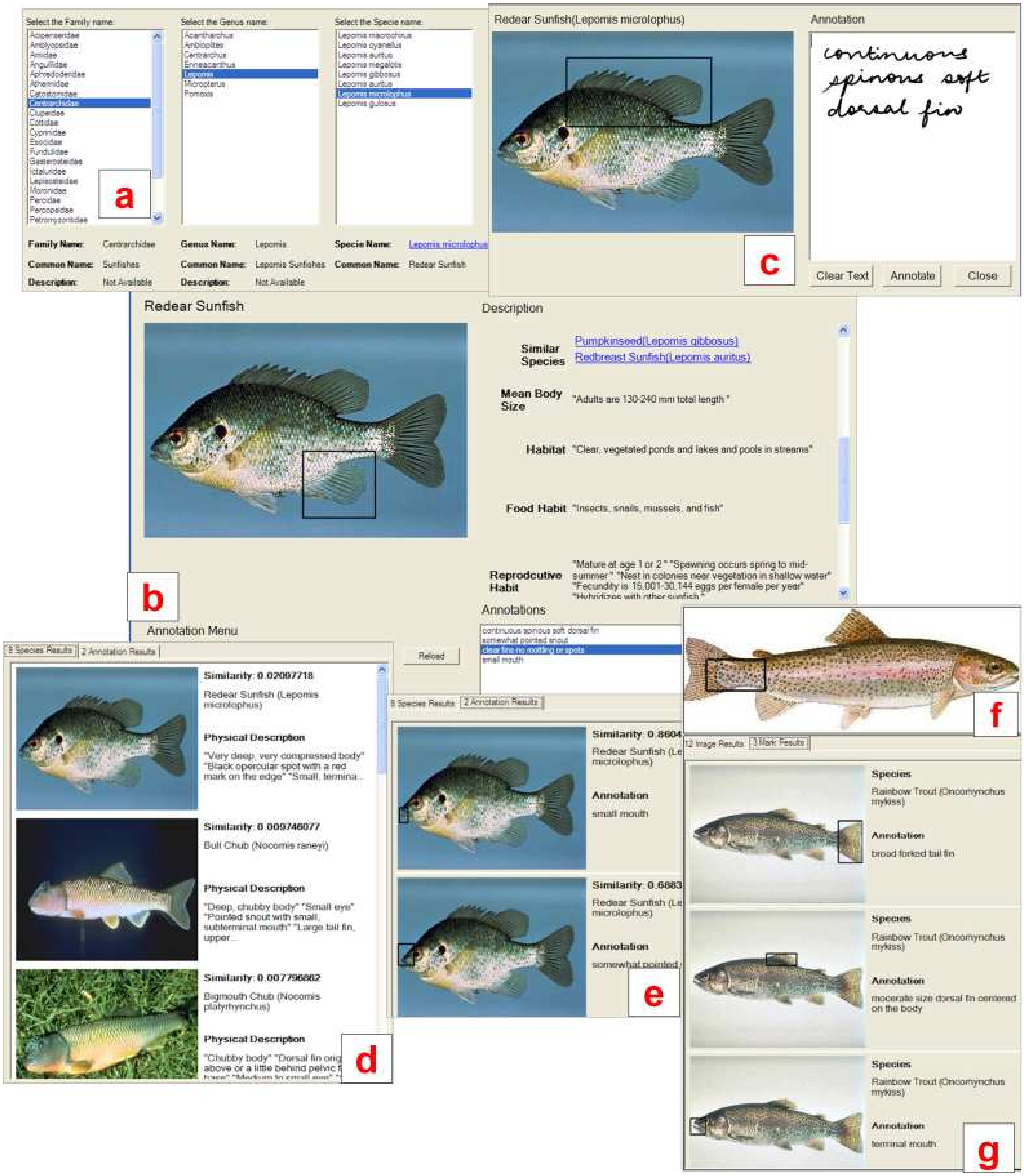

Fig. 14 Screenshots of SuperIDR features: a) Taxonomy browser; b) Species description screen shows details of species and annotations - the highlighted annotation (bottom right) is associated with a marked region in the image; c) Annotation screen – pen input is used to mark the fish image and "write" the annotation, which gets automatically recognized; d) Eight species description results for the text query '`"red mark" "small mouth" "pointed snout" "no spots"`'; e) Two annotation results for the same text query; f) Content-based image search, where the query is the marked region that covers black dots on the body of a rainbow trout; g) Image search results, which can be annotated image marks (shown in the figure) and/or complete images.

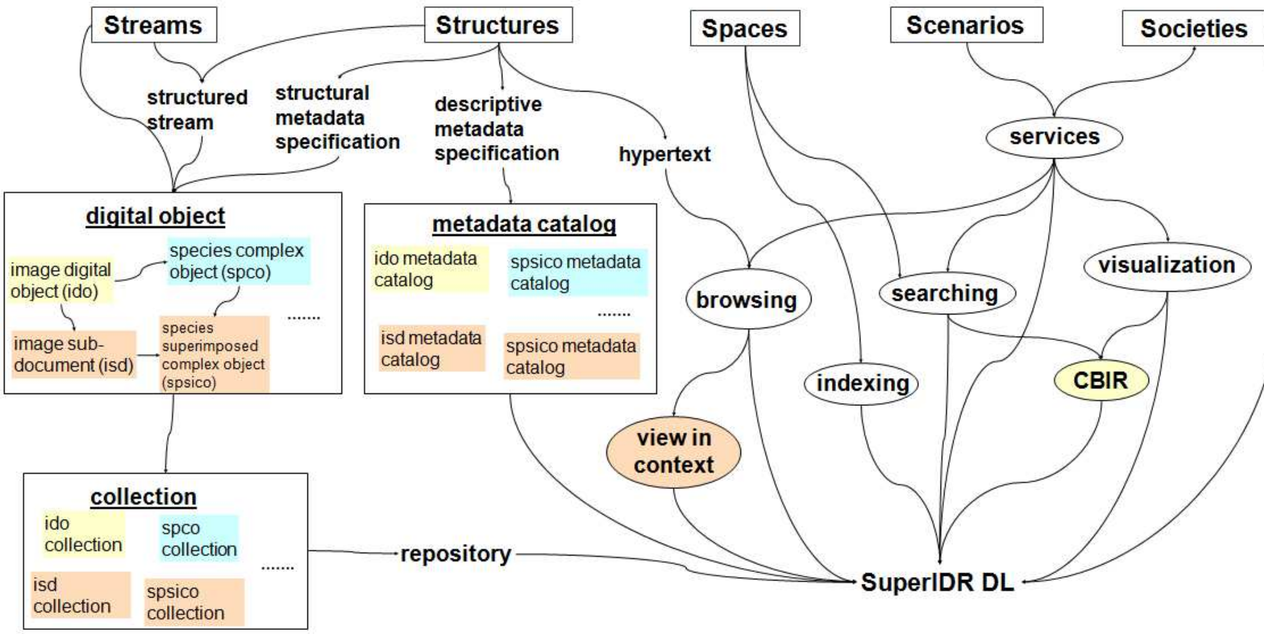

Fig. 16 Definitional dependencies among concepts in an SuperIDR digital library, showing connections among concepts in the 5S framework and the extensions defined.

Each type of digital object is also part of a collection of the same type. We the following notation for each of these types of digital objects:

1. image digital object - *ido*
2. image subdocument - *isd*
3. annotation - *ann*
4. species description - *desc*
5. species complex object - *spco*
6. species superimposed complex object - *spsico*
7. base document - *bd*

Note that each of the first five aforementioned digital objects is a candidate base document. When a subdocument is created on an image digital object, the image digital object becomes a base document in addition to being an image digital object. Thus, one digital object can be part of one or more collections.

We can define a SuperIDR digital library as 4-tuple, $SuperIDR\_DL$ = $(SuperIDR\_\mathscr{R}, SuperIDR\_DM, SuperIDR\_Serv, SuperIDR\_Soc)$, where

– $SuperIDR\_\mathscr{R}$ is a repository, having collections C_{ido}, C_{isd}, C_{ann}, C_{desc}, C_{spco}, C_{spsico}, and C_{bd}, where
  – C_{ido} is a collection of image digital objects,
  – C_{isd} is a collection of image subdocuments,
  – C_{ann} is a collection of annotations
  – C_{desc} is a collection of species descriptions,
  – C_{spco} is a collection of species complex objects,
  – C_{spsico} is a collection of species superimposed complex objects,
  – C_{bd} is a collection of base documents,
– $SuperIDR\_DM = \{DM_{ido}, DM_{isd}, DM_{ann}, DM_{desc}, DM_{spco}, DM_{spsico}, DM_{bd}\}$ is a set of descriptive metadata specifications, where
  – DM_{ido} is a metadata catalog for the collection of image digital objects,
  – DM_{isd} is a metadata catalog for the collection of image subdocuments,
  – DM_{ann} is a metadata catalog for the collection of annotations,
  – DM_{desc} is a metadata catalog for the collection of species descriptions,
  – DM_{spco} is a metadata catalog for the collection of species complex objects,
  – DM_{spsico} is metadata catalog for the a collection of species superimposed complex objects,
  – DM_{bd} is a metadata catalog for the collection of base documents,
– $SuperIDR\_Serv$ is a set of services containing services for indexing, searching, browsing, CBIR and view in context;
– $SuperIDR\_Soc$ of $SuperIDR\_DL$ is a society including {Patron, Fisheries_Student, Fisheries_Faculty, Fisheries_Researchers, SuperIDR_Admin, ... }.

We now describe the contents of some of these components further. The set of streams in $SuperIDR\_DL$ consists of image and text streams. The union set of handles of various digital objects in collections C_{ido}, C_{isd}, C_{ann}, C_{desc}, C_{spco}, C_{spsico}, and C_{bd} will compose $SuperIDR\_DL_{IDs}$, the set of

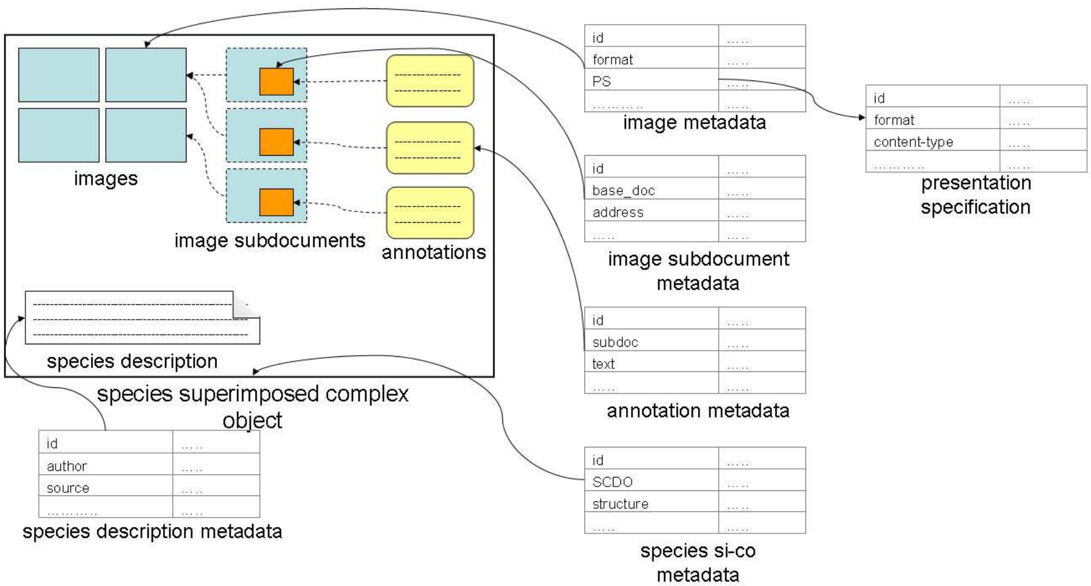

Fig. 17 A species superimposed complex object, its components, associated metadata, and relationships among all of the above.

handles in *SuperIDR_DL*. Examples of content of each metadata specification are described here.

1. DM_{ido} ={'id', 'image name', 'format', 'size', location, ...};
2. DM_{desc} = {'id', 'author', 'source', ...};
3. DM_{spco} = {'id', 'author', 'structure', ... };
4. DM_{bd} = {'id', 'name', 'format', 'size', ...}.
5. DM_{isd} = {'id', 'base document', 'address', 'presentation_specification', ...};
6. DM_{ann} = {'id', 'subdocument', 'text',...};
7. DM_{spsico} = {'id', 'author', 'structure', ... }.

Items 4, 5, and 6 are added to *SuperIDR_DL*, when at least one of the images within the species complex object is marked and annotated. Then, the species complex object is modified into a species superimposed complex object as it now contains subdocuments.

Using *SuperIDR_DL*, we will formally describe three scenarios, each of which involves one or more services of the extensions mentioned in this paper.

1. *AddImageSubdocumentAndAnnotation*
Informal description: This scenario is part of creating and adding an annotation into DLSuperIDR. We focus on what happens in a DLSuperIDR before, during, and after a subdocument is created. Given an image, which is associated with a species, an address referencing a part of the image, and an associated text annotation, a subdocument and an annotation object are created. In addition, the newly created subdocument and annotation are added to the species complex object. If this is the first subdocument added to a species, it changes from being a species complex object to a species superimposed complex object.
Goal: Given an image, which is part of a species complex object, an address of a part of that image, and an associated text annotation, create a subdocument and annotation object and add those to the aforementioned species complex object. This adds a new subdocument to the *DLSuperIDR* and makes the species complex object a species superimposed complex object.
Scenario:
$\langle e_1 : p = AddImageSubdocumentAndAnnotation$ $(ido_j, spco_i, ps_k, addr_l, ann_m), e_2{:}p = response\ (spsico_i, isd_o)\rangle$, where the following constraints apply:
(a) ido_j is an image digital object, such that $ido_j \in spco_i$ and $ido_j \in C_{ido}$ and $spco_i \in C_{spco}$, where $spco_i$ is a species complex object that consists of images and species descriptions, C_{ido} is a collection of image digital objects in *SuperIDR_DL*, and C_{spco} is a collection of species complex objects in *SuperIDR_DL*.
(b) $addr_l$ is an address, specifying a region/span within the image digital object ido_j, and is associated with a presentation specification ps_k.
(c) ann_m is an annotation digital object, such that $ann_m \in spsico_i$' and $ann_m \in C_{ann}$, where C_{ann} is a collection of annotations in *SuperIDR_DL*.
(d) isd_o is a newly created subdocument, such that $isd_o \in spsico_i$' and $isd_o \in C_{isd}$, where where C_{isd} is a collection of image subdocuments in *SuperIDR_DL*.
(e) $spco_i$ is modified into $spsico_i$', a species superimposed complex object, such that ido_j and other digital objects in $spco_i$ are now in $spsico_i$',and $spsico_i$'

$\in C_s psico$, where $C_s psico$ is a collection of species superimposed complex objects in *SuperIDR_DL*.

(f) $C_{spco}' = C_{spco} - spco_i$, where C_{spco}' is the modified collection of species complex objects in *SuperIDR_DL*, which does not contain the species complex object, $spco_i$.

2. *GetImagesAndPartsOfImages*
Informal description: Given an image or a part of image as a query, return a list of images and/or parts of images that match the query image (see Figure 14-f, g). Each image or part of image in the result list also displays other associated information, such as the species description and the annotation text.
Goal: Given an image or a part of an image as query, a set of matching images or parts of images as results.
Scenario: $\langle e_1 : p = GetImages(sm_{sd}[i,j], h_D, k), \rangle$, where
   - $GetImages \in CBISC\_Searching$ is a service and *CBISC_Searching* is a search service in *Super_Serv*. *GetImages* is defined as follows: $\langle e_1 : p = OP_s(Q, ImgC)\rangle$, where $Q = \{q\}$ is a query specification (defined by HTTP request parameters) $q = (H_q, Contents_q, P_q)$, $H_q = ((V_q, E_q), L_q, F_q)$, $V_q = \{v_1, v_2, v_3\}$, $P_q(v_1) = sm_{sd}[i,j]$ (input query image or part of image), $P_q(v_2) = h$, h is an image descriptor handle, $P_q(v_3) = k$ if q is a *KNNQ* or $P_q(v_3) = r$ if q is a *RQ*, $ImgC \in \mathscr{R}$. The computation of OP_s relies on the use of image descriptor $\hat{D} = (h, \varepsilon_D, \delta_D) \in ImgC(2)$ defined by handle $P_q(v_2)$. ε_D is used to extract a feature vector fv_q from $ido_q \in Contents_q$, while δ_D is used to compute the similarity between fv_q and all feature vectors $fv_i \in ICD(1)$, where $ICD \in ido_i(5)$, and $ido_i \in ImgC(1)$.
   - $sm_{sd}[i,j]$ is an image or part of image;
   - h_D is handle of an image descriptor;
   - k is the number of images or part of images to be returned;
   - $p \in C_{ido} \cup C_{isd}$, is composed by images and image subdocuments.

3. *DisplayImageSubdocumentList*
Informal description: This scenario can take place in case of browsing search results (see Figure 14-e, g) which include parts of images and/or browsing through annotations (see Figure 14-b) associated with an image within a species. Given a list of image subdocuments, return a view of the list, clicking of a result item will cause the system to display the subdocument in its context or the context of its containing base document. In a sense a link is being traversed from the subdocument in the list to the subdocument in its original context.
Goal: Given a list of image subdocuments, display them in context of the original base document.
Scenario: $\langle e_1 : p =$ *DisplayImageSubdocumentList*$(isd_1, isd_2, \ldots, isd_n)$,
$e_2 : p = response\,(\mathscr{P}(v_{t_{isd_1}}), \mathscr{P}(v_{t_{isd_2}}), \ldots, \mathscr{P}(v_{t_{isdn}}))$, such that $\mathscr{P}(v_{t_{isd_i}}), 1 \leq i \leq n$ is the response to the service $ViewInContext(v_{k_{isd_i}}, e_{k_i})$, with the following constraints:
(a) $isd_i, 1 \leq i \leq n$ are image subdocuments
(b) $e_{k_i} = (v_{k_{isd_i}}, v_{t_{isd_i}}) \in E_{H_E}$, where E_{H_E} is the extended hypertext formed by the network of image base documents, image subdocuments, and species superimposed complex objects.
(c) $v_{k_{isd_i}}$ is a reference of the image subdocument in the species superimposed complex object
(d) $v_{t_{isd_i}}$ is the subdocument in its original context in its associated image digital object

7 Conclusions and Future Work

Many digital library implementations and applications demand additional and advanced services beyond those found in conventional digital library. Examples of commonly required services include those related to the support of newer, more complex media types such as images, multimedia objects, subdocuments within other documents, or annotations.

In this paper, we address formal definitions and descriptions of desired functionality for DLs by extending the 5S formalism in three areas: content-based image services, complex object services, and superimposed information services. This formalism can help to understand these concepts under the digital library perspective. The set of definitions also may impact future development efforts of a wide range of digital library experts since it can guide the design and implementation of new digital library services based on complex objects, superimposed information, and image content. The proposed concepts were illustrated through the description of case studies as well as potential scenarios that take advantage of complex objects, superimposed information, and content-based image retrieval services.

Future work will include the formalization of more complex services that can be constructed by using the proposed constructs. Examples include multimodal search services, recommendation systems for complex objects and superimposed information, image browsing services based on image content similarity, and management of complex simulation-based content. We also plan to use the proposed formalism to integrate the management of complex objects, superimposed information, and image content descriptions into existing digital library design and implementation tools [28].

Acknowledgements We thank our financial sponsors of current and past projects, including NSF (Ensemble - DUE 0840713, Sidewalk Project - DUE 0435059), Microsoft corporation (SuperIDR Project - Microsoft Tablet PC grant), HP Technology for Teaching, CAPES, FAPESP, and the CNPq (BioCORE Project). We also thank F. Corubolo, L. Cassel, S. Price, D. Maier, S. Murthy, J. Mccall, K. Kuchibhotla, and R. Suryavanshi. We are grateful to the American Fisheries Society for

use of copyrighted material from Jenkins and Burkheads Freshwater Fishes of Virginia.

References

1. (2000) Text encoding initiative: Yesterday's information tomorrow. URL http://www.tei-c.org/
2. (2001) Bundles in captivity: An application of superimposed information. Tech. rep., Washington, DC, USA
3. (2003) Dspace. URL http://www.dspace.org/
4. (2003) Fedora repository. URL http://www.fedora-commons.org/
5. (2009) Flickr - photo sharing. URL http://www.flickr.com/
6. Agosti M, Ferro N (2008) A formal model of annotations of digital content. Transactions on Information Systems (TOIS) 26(1):1–55
7. Agosti M, Ferro N, Frommholz I, Thiel U (2004) Research and Advanced Technology for Digital Libraries, chap Annotations in Digital Libraries and Collaboratories - Facets, Models and Usage, pp 244–255. URL http://www.springerlink.com/content/ah229da5je3ka8y7
8. Agosti M, Albrechtsen H, Ferro N, Frommholz I, Hansen P, Orio N, Panizzi E, Pejtersen AM, Thiel U (2005) DiLAS: a digital library annotation service
9. Alavi M, Leidner DE (2001) Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. MIS Quarterly 25(1):107–136, URL http://www.jstor.org/stable/3250961
10. Archer DW, Delcambre LM, Corubolo F, Cassel L, Price S, Murthy U, Maier D, Fox EA, Murthy S, Mccall J, Kuchibhotla K, Suryavanshi R (2008) Superimposed information architecture for digital libraries. In: ECDL '08: Proceedings of the 12th European conference on Research and Advanced Technology for Digital Libraries, Springer-Verlag, Berlin, Heidelberg, pp 88–99, DOI http://dx.doi.org/10.1007/978-3-540-87599-4_10
11. Arko RA, Ginger KM, Kastens KA, Weatherley J (2006) Using annotations to add value to a digital library for education. D-Lib Magazine 12(5), URL http://www.dlib.org/dlib/may06/arko/05arko.html
12. Atnafu S, Chbeir R, Coquil D, Brunie L (2004) Integrating similarity-based queries in image dbmss. In: Proceedings of the 2004 ACM symposium on applied computing, pp 735–739
13. Bargeron D, Brush BAJ, Gupta A (2001) A common annotation framework. Tech. rep., Microsoft Corporation: MSR-TR-2001-108, URL http://research.microsoft.com/apps/pubs/default.aspxid=69900
14. Barrett C, Bisset K, Eubank S, Feng X, Marathe M (2008) Episimdemics: an efficient and scalable framework for simulating the spread of infectious disease on large social networks. ACM/IEEE Conference on Supercomputing
15. Barrett C, Bisset K, Leidig J, Marathe A, Marathe M (2009) Estimating the impact of public and private strategies for controlling an epidemic: A multi-agent approach. 21st Innovative Applications of Artificial Intelligence
16. Bergman LD, Castelli V, Li CS (1997) Progressive content-based retrieval from satellite image archives. D-Lib Magazine 3(10)
17. Bolettieri P, Falchi F, Gennaro C, Rabitti F (2007) Automatic metadata extraction and indexing for reusing e-learning multimedia objects. In: MS '07: Workshop on multimedia information retrieval on The many faces of multimedia semantics, ACM, New York, NY, USA, pp 21–28, DOI http://doi.acm.org/10.1145/1290067.1290072
18. Burnett IS, Pereira F, de Walle RV, , Koenen R (2006) The MPEG-21 Book. John Wiley & Sons
19. Candela L, Castelli D, Ferro N, Ioannidis Y, Koutrika G, Meghini C, Pagano P, Ross S, Soergel D, Agosti M, Dobreva M, Katifori V, Schuldt H (2008) The delos digital library reference model - foundations for digital libraries. version 0.98. DELOS Network of Excellence on Digital Libraries
20. Castelli D, Ioannidis Y, Ross S, Schek HJ, Schuldt H (2006) DELOS Network of Excellence on Digital Libraries - Reference Model for DLMS
21. Ciaccia P, Patella M, Zezula P (1997) M-tree: An Efficient Access Method for Similarity Search in Metric Spaces. In: Proceedings of 23rd International Conference on Very Large Data Bases, Athens, Greece, pp 426–435
22. Delcambre LM, Archer D, Price S, Murthy U, Fox EA, Cassel L (2008) Superimposing a strand map over a database lecture. Presented at the 39th SIGCSE technical symposium on Computer science education
23. Delcambre LML, Maier D (1999) Models for superimposed information. In: ER '99: Proceedings of the Workshops on Evolution and Change in Data Management, Reverse Engineering in Information Systems, and the World Wide Web and Conceptual Modeling, Springer-Verlag, London, UK, pp 264–280
24. Dourado P, Ferreira P, Santanchè A (2006) Representação unificada de objetos digitais complexos: Confrontando o ras com o ims cp. In: III Workshop de Bibliotecas Digitais.
25. Federation DL (2009) Metadata encoding and transmission standard (mets). URL http://www.loc.gov/standards/mets/
26. Flickner M, Sawhney H, Niblack W, Ashley J, Huang Q, Dom B, Gorkani M, Hafner J, Lee D, Petkovic D, Steele D, Yanker P (1995) Query by image and video

content: the qbic system. Computer 28(9):23–32

27. French JC, Chapin AC, Martin WN (2003) An application of multiple viewpoints to content-based image retrieval. In: Proceedings of the 3rd ACM/IEEE-CS joint conference on digital libraries, IEEE Computer Society, Washington, DC, USA, pp 128–130
28. Goncalves MA (2004) Streams, Structures, Spaces, Scenarios, and Societies (5S): A Formal Digital Library Framework and Its Applications. PhD thesis, Virginia Tech, Blacksburg, VA
29. Goncalves MA, Fox EA (2002) 5SL: a language for declarative specification and generation of digital libraries. In: Proceeding of the Second ACM/IEEE-CS Joint Conference on digital libraries, Portland, Oregon, USA, pp 263–272
30. Gonçalves MA (2004) Streams, structures, spaces,scenarios, and societies (5S): A formal digital library framework and its applications. PhD thesis, Virginia Tech, Blacksburg
31. Gonçalves MA, Fox EA, Watson LT, Kipp NA (2004) Streams, structures, spaces, scenarios, societies (5S): A formal model for digital libraries. ACM Trans Inf Syst 22(2):270–312
32. Gony J, Cord M, Philipp-Foliguet S, Gosselin PH, Precioso F, Jordan M (2007) Retin: a smart interactive digital media retrieval system. In: CIVR ’07: Proceedings of the 6th ACM international conference on Image and video retrieval, ACM, New York, NY, USA, pp 93–96, DOI http://doi.acm.org/10.1145/1282280.1282294
33. Halasz F, Schwartz M (1994) The Dexter hypertext reference model. Commun ACM 37(2):30–39
34. Hardman L, Bulterman DCA, van Rossum G (1994) The Amsterdam hypermedia model: adding time and context to the Dexter model. Commun ACM 37(2):50–62, URL http://dx.doi.org/10.1145/175235.175239
35. Hong JS, Chen H, Hsiang J (2000) A Digital Museum of Taiwanese Butterflies. In: Proceedings of the Fifth ACM Conference on digital Libraries, San Antonio, Texas, United States, pp 260–261
36. Jenkins RE, Burkhead NM (1994) Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, MD
37. Jin H, Sun A, Zheng R, He R, Zhang Q, Shi Y, Yang W (2007) Content and semantic context based image retrieval for medical image grid. In: GRID, pp 105–112
38. Kahan J, Koivunen MR (2001) Annotea: an open RDF infrastructure for shared web annotations. In: WWW ’01: Proceedings of the 10th international conference on World Wide Web, ACM, New York, NY, USA, pp 623–632, DOI http://doi.acm.org/10.1145/371920.372166
39. Kerne A (2005) Recombinant Information Workshop [interface ecology lab]
40. Kerne A, Koh E, Dworaczyk B, Mistrot JM, Choi H, Smith SM, Graeber R, Caruso D, Webb A, Hill R, Albea J (2006) combinformation: a mixed-initiative system for representing collections as compositions of image and text surrogates. In: JCDL 2006: Proceedings of the 6th ACMIEEE-CS joint conference on Digital libraries, pp 11–20
41. Krafft DB, Birkland A, Cramer E (2008) Ncore: architecture and implementation of a flexible, collaborative digital library. In: JCDL ’08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, ACM, New York, NY, USA, pp 313–322, DOI http://doi.acm.org/10.1145/1378889.1378943
42. Lagoze C, de Sompel HV (2001) The open archives initiative: Building a low-barrier interoperability framework. In: JCDL ‘01: In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, ACM Press, pp 54–62
43. Lagoze C, de Sompel HV (2001) The Open Archives Initiative: Building a Low-Barrier Interoperability Framework. In: Proceedings of the Joint Conference on Digital Libraries, Roanoke, VA, USA, pp 54–62
44. Lagoze C, Sompel HV (2007.) URL http://www.omg.org/docs/formal/05-11-02.pdf
45. Learning AD (2004) Sharable content object reference model (scorm). URL http://www.adlnet.org/index.cfm?fuseaction = SCORDown
46. Lynch C, Parastatidis S, Jacobs N, de Sompel HV, Lagoze C (2007) The oai-ore effort: progress, challenges, synergies. In: JCDL ’07: Proceedings of the 2007 conference on Digital libraries, ACM, New York, NY, USA, p 80, DOI 10.1145/1255175.1255190, URL http://dx.doi.org/10.1145/1255175.1255190
47. Lyons J, Hanson P, White E (2006) A photo-based computer system for identifying wisconsin fishes. Fisheries 31(6):269–275
48. Maier D, Delcambre LML (1999) Superimposed information for the internet. In: WebDB (Informal Proceedings), pp 1–9, URL http://citeseer.ist.psu.edu/258572.html
49. Marshall CC (1997) Annotation: from paper books to the digital library. In: DL ’97: Proceedings of the second ACM international conference on Digital libraries, ACM Press, New York, NY, USA, pp 131–140, URL http://portal.acm.org/citation.cfm?id=263806
50. Melton J, Eisenberg A (2001) Sql multimedia and application packages (sql/mm). SIGMOD Rec 30(4):97–102, DOI http://doi.acm.org/10.1145/604264.604280
51. Murthy S, Maier D, Delcambre L (2004) Querying bi-level information. In: Proceedings of WebDB workshop, URL http://webdb2004.cs.columbia.edu/papers/1-2.pdf
52. Murthy S, Maier D, Delcambre L, Bowers S (2004) Superimposed applications using SPARCE. In: Proceed-

ings of the 20th International Conference on Data Engineering, IEEE Computer Society, Los Alamitos, CA, USA

53. Murthy S, Delcambre L, Maier D (2006) Explicitly representing superimposed information in a conceptual model. In: Conceptual Modeling - ER 2006, pp 126–139, URL http://dx.doi.org/10.1007/11901181_11
54. Murthy S, Delcambre L, Maier D (2006) Explicitly representing superimposed information in a conceptual model. In: Conceptual Modeling - ER 2006, Springer-Verlag, Berlin, Heidelberg, pp 126–139, DOI http://dx.doi.org/10.1007/11901181_11
55. Murthy U, Ahuja K, Murthy S, Fox EA (2006) SIMPEL: a superimposed multimedia presentation editor and player. In: JCDL '06: Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries, p 377
56. Murthy U, Richardson R, Fox EA, Delcambre L (2006) Enhancing concept mapping tools below and above to facilitate the use of superimposed information. In: Cañas AJ, Novak JD (eds) the Second International Conference on Concept Mapping, URL http://cmc.ihmc.us/cmc2006Papers/cmc2006-p197.pdf
57. Murthy U, da Silva Torres R, Fox EA (2006) Sierra - a superimposed application for enhanced image description and retrieval. In: ECDL'06: Proceedings of the 10th European conference on Research and Advanced Technology for Digital Libraries, Springer-Verlag, vol 4172, pp 540–543
58. Murthy U, Gorton D, Torres R, Gonçalves M, Fox E, Delcambre L (2007) Extending the 5S digital library (DL) framework: From a minimal DL towards a DL reference model. URL http://si.dlib.vt.edu/publications/2007_DLRL_DLF_v4.pdf, presented at the First Workshop on Digital Library Foundations - held in conjunction with JCDL'07
59. Murthy U, Fox EA, Chen Y, Hallerman E, Torres R, Ramos E, Falcao T (2009) Superimposed image description and retrieval for fish species identification. 13th European Conference on Research and Advanced Technology for Digital Libraries
60. Murthy U, Fox EA, Chen Y, Hallerman E, Torres R, Ramos EJ, Falcao TR (2009) Species identification: fish images with cbir and annotations. In: JCDL '09: Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries, ACM, New York, NY, USA, pp 435–436, DOI http://doi.acm.org/10.1145/1555400.1555500
61. Nelson L, de Sompel HV (2006) Ijdl special issue on complex digital objects: Guest editors' introduction. International Journal of Digital Libraries 6(2):113–114, DOI http://dx.doi.org/10.1007/s00799-005-0127-y.
62. Nelson ML, Argue B, Efron M, Denn S, Pattuelli MC (2001) A survey of complex object technologies for digital libraries. Tech. rep.
63. Nelson TH (1999) Xanalogical structure, needed now more than ever: parallel documents, deep links to content, deep versioning, and deep re-use. ACM Comput Surv 31(4es), URL http://portal.acm.org/citation.cfm?id=346033
64. Pahl C, Barrett R (2004) A web services architecture for learning object discovery and assembly. In: WWW Alt. '04: Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters, ACM, New York, NY, USA, pp 446–447, DOI http://doi.acm.org/10.1145/1013367.1013518
65. Persoon E, Fu K (1977) Shape Discrimination Using Fourier Descriptors. IEEE Transanctions on Systems, Man, and Cybernetics 7(3):170– 178
66. Ravindranathan U, Shen R, Gonçalves MA, Fan W, Fox EA, Flanagan JW (2004) ETANA-DL: managing complex information applications – an archaeology digital library. In: JCDL '04: Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM Press, New York, NY, USA, pp 414–414
67. Rehberger D, Fegan M, Kornbluh M (2006) Research and Advanced Technology for Digital Libraries, chap Reevaluating Access and Preservation Through Secondary Repositories: Needs, Promises, and Challenges, pp 39–50
68. Shen R, Vemuri NS, Fan W, da S Torres R, Fox EA (2006) Exploring digital libraries: integrating browsing, searching, and visualization. In: JCDL '06: Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM, New York, NY, USA, pp 1–10, DOI http://doi.acm.org/10.1145/1141753.1141755
69. for Space Data Systems CC (2002) Reference model for an open archival information system (oais) : Recommendation for space data system standards : Ccsds 650.0-b-1. Tech. rep.
70. Suleman H, Fox EA (2002) Beyound harvesting: Digital library components as oai extensions. technical report.
71. Swain M, Ballard D (1991) Color Indexing. International Journal of Computer Vision 7(1):11–32
72. Thompson HS, Connolly D (1997-2003) W3c xml pointer, xml base and xml linking. URL http://www.w3.org/XML/Linking, description of work done by the W3C XML Linking Working Group
73. Torres RS (2004) Integrating image and spatial data for biodiversity information management. PhD thesis, Institute of Computing, University of Campinas
74. Torres RS, ao AXF (2006) Content-Based Image Retrieval: Theory and Applications. Revista de Informática Teórica e Aplicada 13(2):161–185
75. Torres RS, Falcão AX, da F Costa L (2004) A Graph-based Approach for Multiscale Shape Analysis. Pattern Recognition 37(6):1163–1174

76. Torres RS, Medeiros CB, Goncalves MA, Fox EA (2006) A digital library framework for biodiversity information systems. International Journal on Digital Libraries 6(1):3–17
77. Traina C, Figueiredo JM, Traina AJM (2005) Image domain formalization for content-based image retrieval. In: Proceedings of the 2005 ACM symposium on applied computing, pp 604–609
78. Veltkamp RC, Tanase M (2002) Content-based image retrieval systems: A survey. Technical Report TR UU-CS-2000-34, Department of Computing Science, Utrecht University
79. Wang JZ, Du Y (2001) Scalable integrated region-based image retrieval using irm and statistical clustering. In: Proceedings of the 1st ACM/IEEE-CS joint conference on digital libraries, pp 268–277
80. Wang Y, Makedon F, Ford J, Shen L, Goldin D (2004) Generating fuzzy semantic metadata describing spatial relations from images using the r-histogram. In: Proceedings of the 4th ACM/IEEE-CS joint conference on digital libraries, pp 202–211
81. Winget M (2007) Annotation functionality for digital libraries supporting collaborative performance: an example of musical scores. In: JCDL '07: Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries, ACM, New York, NY, USA, pp 305–306, DOI http://doi.acm.org/10.1145/1255175.1255234
82. Zhu B, Ramsey M, Chen H (2000) Creating a Large-Scale Content-Based Airphoto Image Digital Library. IEEE Transactions on Image Processing 9(1):163–167

Appendix

Table 3 Detailed Scenario of Table 1.

1	Content-based image retrieval (CBIR)	Brad is a PhD student in Veterinary Medicine. He has been working on a research to find the effect of diabetes on a mouse's fetus, which involves identification and comparison of multiple images of the fetus's dissected heart. To find all the similar images, Brad specifies a query using a fetus's heart image that shows mutated part. The system returns a list of matching images ranked according to similarity. From the list, he selects several images and could be able to compare them side-by-side for details.
2	Complex objects	Rahul is doing research on parasites. His current task is to group images of species as well as related information and store them into a digital library. He has multiple images of Eurytrema pancreaticum, which is found in cattle and buffalos. Each image has different zoom level and resolutions and there is much information to be stored along with the images such as its family, subfamily, genus, species, habitat, hosts, etc. A means to store all those information as a unified group is used to enhance handling as well as effective search/browse services.
3	Superimposed information and services	A Computer Science professor is preparing for a class on the system simulation. Most of her class material comes from existing multimedia as well as text resources. She wants to work with pieces of information in various documents highlighting it in the context of original resource. She selects a portion of a well-known paper and types in her annotation and does this on an image of a simulation model diagram, too. This is stored into a digital library, where her students access to view the original paper and diagrams along with her annotations. The portion that she selected for annotation is highlighted to direct the viewer's attention.
4	Integrated Service	Jason is a researcher at the Institute of Biological Simulations. He specializes in the Epidemiological simulation using a high-performance computer, which visually shows patterns of how an epidemic spreads in a population. It requires multi-disciplinary knowledge such as Biology, Geographic Information Science and Social Science to understand statistical models and factors involved to create a computer simulation. He captures a screenshot during a simulation then selects a part that shows a unique pattern and annotates on it. He also links the selected part to Web resources. To find relevant patterns, he specifies this unique pattern image to a software tool, which searches the digital library for similar shape and texture characteristics. A list of result is returned ranked by its similarity. He browses through the list and links one pattern to the simulation screenshot. All of these annotations, screenshots of simulation, input parameters and simulation results along with tags and links to other images are stored in a digital library as a unified group of information for sharing, reference and preservation purposes.
5	Integrated Service	Matt is majoring in Fisheries, enrolled in the Ichthyology class. In the past, he has supplemented the use of dichotomous keys with personal notes, pictures from the Web and textbooks. From this semester, he has been using a software tool, which runs on a tablet PC. Using it, he browses to an image of a red-ear sunfish to see the physical description, habitat, food habits, etc. He then adds an annotation using a pen input on the image by selecting and associating a portion of it with notes and then he links it with Web resources. All of this information entered by him as well as original image and metadata are stored in a digital library as a unified group of information. In the field, Matt is examining an unknown fish specimen that he collected. He takes a picture of it and enters the picture as a query to the software tool. The tool matches the image with other similar images stored in a digital library and returns results ranked by their similarity. Matt selects one on top and finds information such as species, description, habitat, etc.