Digital libraries and Web 3.0. The CallimachusDL approach

Ángel García-Crespo [a,1], Juan Miguel Gómez-Berbís [a], Ricardo Colomo-Palacios [a,*],
Francisco García-Sánchez [b,2]

[a] Computer Science Department, Universidad Carlos III de Madrid, Av. Universidad 30, Leganés 28911, Madrid, Spain
[b] Computer Science Department, Universidad de Murcia, Campus de Espinardo, 30071 Espinardo, Murcia, Spain

ARTICLE INFO

ABSTRACT

The constantly increasing volume of information available on the Internet is changing the forms of clas sification and access to data. Given the immense collection of information stored on the Internet, digital libraries constitute a fundamental subject of research. Among the challenges of classifying, locating and accessing knowledge in digital libraries tackling with the huge amount of resources the Web provides, improving digital libraries by means of different strategies, particularly, using semantics remains a prom ising and interesting approach. In this paper, CallimachusDL is presented, a semantics based digital library which provides faceted search, enhanced access possibilities and a proof of concept implementa tion. CallimachusDL represents a novel approach to digital libraries, integrating social web and multime dia elements in a semantically annotated repository. The results of the implementation indicate that the features proposed in CallimachusDL are encouraging and extendable in the use of digital libraries.

1. Introduction

Digital libraries represent a new breed of software applications whose aim encompasses categorizing, classifying, archiving and providing access to the vast constellation of Web resources. Cur rently, digital libraries (DL for short) are facing a new paradigm shift coping with various challenges which include overcoming tra ditional browsing or keyword based strategies. Fundamentally, DL infrastructure improvement attempts have been trying to increase the quality of information retrieval, from query expansion to col laborative filtering or multi faceted browsing (Kruk & Decker, 2007; Lytras, Damiani, & Ordóñez de Pablos, 2008; Lytras & Ordóñ ez de Pablos, 2009). However, current approaches are still not ful filling expectations, leading the user in many cases to frustration.

On the other hand, the Semantic Web has emerged to be a new and highly promising context for knowledge and data engineering (Vossen, Lytras, & Koudas, 2007). The term "Semantic Web" was coined by Berners Lee, Hendler, and Lassila (2001), to describe the evolution from a document based web towards a new para digm that includes data and information for computers to manip ulate. The Semantic Web enables automated information access based on machine processable semantics of data. This means that this data will be available for providing precise and exhaustive information retrieval. Thus, the Semantic Web provides a comple mentary vision as a knowledge management environment (War ren, 2006) that, in many cases has expanded and replaced previous knowledge and information management archetypes (Da vies, Lytras, & Sheth, 2007; Lytras & García, 2008; Lytras & Ordóñez de Pablos, 2007, 2009; Ordóñez de Pablos, 2002; Rodriguez Pérez & Ordóñez de Pablos, 2003). Semantic Web has been named as Web 3.0 (Hendler, 2008; Lassila & Hendler, 2007) as a new version of Web 2.0 in which web has advanced to become what Tim Bern ers Lee (2007) has termed the "Giant Global Graph".

According to Lytras and García (2008), in recent years, Semantic Web research has resulted in significant outcomes and the adop tion of this technology from the market and the industry is becom ing closer. Thus, the application of semantics to knowledge management is not new (e.g. Colomo Palacios, Gómez Berbís, García Crespo, & Puebla Sánchez, 2008; García Crespo, Colomo Palacios, Gómez Berbís, & García Sánchez, 2010, García Crespo, Colomo Palacios, Gómez Berbís, & Mencke, 2009; García Crespo, Colomo Palacios, Gómez Berbís, & Ruiz Mezcua, 2010). Ontologies (Fensel, 2002) are the technological cornerstones of the Semantic Web, because they provide structured vocabularies that describe a formal specification of a shared conceptualization. Ontologies were developed in the field of Artificial Intelligence to facilitate knowledge sharing and reuse (Fensel et al., 2001). Ontologies pro vide a common vocabulary for a domain and define, with different levels of formality, the meaning of the terms and the relations between them. Knowledge in ontologies is mainly formalized using five kinds of components: classes, relations, functions, axioms and

---

* Corresponding author. Tel.: +34 91 624 5958; fax: +34 91 624 9129.
  E-mail addresses: angel.garcia@uc3m.es (Á. García-Crespo), juanmiguel.gomez@uc3m.es (J.M. Gómez-Berbís), ricardo.colomo@uc3m.es (R. Colomo-Palacios), frgarcia@um.es (F. García-Sánchez).
[1] Tel.: +34 91 624 9417; fax: +34 91 624 9129.
[2] Tel.: +34 968 39 8107; fax: +34 968 36 4151.

instances (Gruber, 1993). Ontologies support the necessary seman
tics to describe text strings (Gruber, 1993; Studer, Benjamins, &
Fensel, 1998).

The domain of DL has not been ignored during the growth of the
Semantic Web. The use of semantics in DL can outperform the cur
rent endeavors that require finding data spread out across the DL
structure and dynamically drawing inferences, something continu
ally hampered by their reliance on ad hoc, task specific frameworks
in present DL technologies. In this paper, CallimachusDL is pre
sented, a semantics based DL which uses semantic information
gathering and browsing to enhance search and retrieval. One of
the originalities of our work is to include social web multimedia
content semantic annotation as a part of the DL.

The remainder of the paper is organized as follows. In Section 2 an
overview of the state of the art is provided. In Section 3, a number of
requirements are discussed and the benefits of tackling them with a
semantically enhanced approach are presented. In Section 4 Calli
machusDL is described in detail, as well as its architecture and
proof of concept implementation. Section 5 provides an experimen
tal setup for CallimachusDL. Finally, Section 6 concludes the paper
providing a number of conclusions and summarizes future work.

2. Digital libraries: been there, done that

DLs provide high quality and well organized information. Many
of the powerful characteristics of DLs rely on Metadata. Librarians
describe the resources of catalogues and other collections through
metadata in order to facilitate efficient delivery of information. The
use of metadata in its formats and functionalities has been an object
of study in the past in the field of DLs, for example the use of XML
(Kim & Choi, 2000) and RDF (e.g. Baruzzo et al., 2009; Bygstad, Ghi
nea, & Klæboe, 2009; Han, 2006; Isaac, Schlobach, Matthezing, &
Zinn, 2008). The use of ontologies in the context of DLs (e.g. Prasad
& Madalli, 2008; Schreiber et al., 2008) could be interesting in order
to incorporate new functionalities by describing the relationships
between elements. The concept of ontology introduced by the
Semantic Web is a promising path to extend DL formalisms with
meaningful annotations (Kruk, Decker, & Zieborak, 2005). Several
authors have proposed ontologies for describing the relationships
between all the elements which comprise a DL scenario (Ferrán,
Mor, & Minguillón, 2005) which go beyond different standards of
DLs description formats such as MARC21, Dublin Core and BibTeX.

The new and promising DLs content management tool genera
tion comes from the combination of the Semantic Web and the
new social aspects of what has been termed the Social Web. Here
several initiatives can be found such as the ambitious JeromeDL
project (Kruk & Decker, 2007) or DLibra (Mazurek & Werla,
2005). JeromeDL uses MarcOnt Ontology (Kruk, Synak, & Zimmer
mann, 2005) which uses several legacy metadata standards
(MARC21, BibTeX and Dublin Core) and offers a number of search
and retrieval services based on semantic technology.

Fundamentally, the CallimachusDL approach is radically differ
ent to the ones detailed before since semantic navigation is pro
posed, along with faceted search and browsing, metadata
representation format and usability as the main building principles
of the entire approach. Those features are detailed in the next
Section.

3. Using semantic information gathering and browsing to enhance search and retrieval

Since the first initiatives for their development, DLs on the Web
had to strive for classifying, locating and accessing resources. How
ever, the advantage of the simplicity in DL leads to their great
drawback, the increasing volume of information being stored

without a clear structure. Actually, most current DL cannot be used
as fully fledged environments to create and search knowledge in
an efficient way, since the information collected in these systems
lies unused by computers, mainly due to the human language in
which the resources are written. As further processing is needed,
new formal approaches are used to make computers "understand"
the Web content or, more precisely, the application of semantics.

The semantic technologies paradigm is based on this statement,
where the traditional Web is enhanced with formal knowledge
placed below the current information. This is possible thanks to
the extensibility of the Web with metadata and metadata process
ing, which allows computational reasoning and intelligent capabil
ities. In the following, the problems which arise when developing a
semantically enhanced DL environment are analyzed, including
technical and social factors:

- *Metadata representation format:* Metadata support for the actual
  information within DL resources must be explicitly declared.
  Some of the current social tools such as the emerging Web 2.0
  applications like Flickr (http://www.flickr.com) or del.icio.us
  (http://del.icio.us/) apply the so called "folksonomies" (a term
  derived from the merging of the words 'folk' and 'taxonomy'.)
  to add meta information in form of tags chosen by the user.
  In this case, tags are different among different users, because
  they are chosen freely, so they cannot be fully exploited in a
  community. However, unfortunately this type of metadata is
  not supported by Semantic Web technology.
- *Navigation:* Ordinary DL base the relationship between pages in
  explicit hyperlinks. These links relate one page to another basi
  cally according to user considerations. If the relation between
  DL resources were represented by means of semantics, the
  application would be able to provide mechanisms to semanti
  cally navigate between related resources with real meaning.
- *Search:* Given a set of resources, the basic type of querying in
  current DL is the keyword based search. Structured requests
  for more advanced information retrieval are needed to make a
  DL a really useful knowledge repository. In addition to simple
  full text searches, users would recover information by querying
  or selecting the semantic knowledge.
- *Usability:* Communities need a critical mass of users. Not only
  the number of users is crucial, but also their participation in
  the communities. The Semantic Web community has to grasp
  this principle and develop it. For that purpose, applications
  enhanced with semantic functionalities have to be designed
  with maximum usability and minimum cognitive load for every
  user, including both Semantic Web experts and Internet users
  with no knowledge about semantics.

The motivation for this work proposes a solution to these require
ments. CallimachusDL is focused on the aforementioned require
ments, solving them, and proposes an integrated solution that
uses semantic information gathering and browsing to enhance
search and retrieval. In the next Section, the approach for Calli
machusDL will be discussed.

4. CallimachusDL: bringing the library mess into order

The current section describes the structure and features of Cal
limachusDL, outlining the resources used which comprise its com
ponents. These include metadata representation format, multi
ontologies, semantic navigation, usability, and faceted search.

4.1. The CallimachusDL description

Given the aforementioned problems that traditional DL cope
with, the CallimachusDL approach is based on several design

principles to avoid these drawbacks, and built as a kernel to devel op a fully fledged semantic working environment for the final users. These design principles are as follows:

- *Metadata representation format:* Bearing in mind that metadata processing requires a controlled and well defined vocabulary, the Semantic Web adopted and developed ontologies as the best mechanism to represent, share and reuse the knowledge hidden within data. An ontology is a formal specification of a shared conceptualization. Since semantic knowledge must be represented in the form of well designed ontologies, the models and language which actually realize the representation of the knowledge must be selected. In order to construct an accurate representation, the different World Wide Web Consortium (W3C) proposed standards are recommended. Resource Description Framework (RDF, http://www.w3.org/TR/rdf pri mer/) and RDF Schema (RDFS, http://www.w3.org/TR/rdf schema/) can be perfectly suitable for defining the semantic information needed. Other languages such as the Web Ontology language (OWL) can also be suitable, but its additional inference mechanism is not required for the real necessities of the application.

- *Multi ontology approach for defining DL resources:* Once the ontologies representation has been defined, the scope of the ontologies used must be explicitly declared. Since DL resources are basically resources on the Web, they should initially be described this way. For this, the Dublin Core initiative (DC, http://dublincore.org/) fits perfectly as the main ontology for describing the wiki pages. Once identified, the DL resources must be described with regard to content. Therefore, a second ontology or several ontologies must be used for formalizing the real domain of the DL resources content.

- *Semantic navigation:* As ordinary hyperlinks are not enough to show the related information in a DL, another approach is needed to offer the user all the information semantically when viewing data. This is mainly due to the fact that the user inter face must enable navigation to semantically related items (Tee van, Alvarado, Ackerman, & Karger, 2004). For that, semantic links, semalinks are proposed in Fuentes Lorenzo, Morato, and Gómez Berbís (2009). Semalinks are ordinary hyperlinks in appearance but built upon semantic information. This semantic information, consisting both of the ontology concept to which a particular part of the content is referring and its value, will lead the user to pages with content semantically similar to the sema link. That is, if a set of words have been used to form a semalink, with a property x and a value, when a mouse is placed over this link, the nodes appearing will make reference to other pages with the same property x and value, and many more references as properties directly related to property x which exist in the repositories, with the same value.

- *Usability:* Authoring a semantic wiki must be made just as easy as authoring a traditional wiki. For that purpose, editing the semantic links must be done at the same time and in the same view as editing the rest of the page. Semantic annotations are the answer to fill this gap. Annotating a document means add ing semantic data to these documents (McEnery & Wilson, 2001). Users will be provided with semantic information to add; therefore, while editing a page, they will be able to anno tate a word or a set of words with semantic data, just as easy as marking the selected words and associating them with a prop erty or vocabulary concept from the ontology domain. Usability is also reflected in the functionalities of browsing and searching seen in the previous subsections.

- *Faceted search:* As keyword based searches or other different syntactical queries are not an efficient retrieval mechanism, and providing that semantic information is underlying the sys tem, a more advanced search is required. A facets based search is the solution. With faceted metadata (Ranganathan, 1962), the information space is partitioned using orthogonal conceptual dimensions of the data. These dimensions are called facets, and represent the characteristics of the information elements. These facets are used then to select or filter the relevant ele ments in a certain information space, leading users to the exact information needed. These facets are the properties defined in the domain ontologies. Technological solution for faceted search is used in previous works by authors (Fuentes Lorenzo et al., 2009).

The main features of CallimachusDL have been described above. In the subsequent Section, CallimachusDL itself is detailed.

4.2. Architecture

The CallimachusDL architecture is heavily based on the SWAN architecture (Fuentes Lorenzo, Gómez Berbís, & García Crespo, 2007). Taking into account these apparently different levels of knowledge (ontologies, resources and semantic information), this knowledge can be divided into three layers:

- *Resource layer:* This layer stores the DL resources and all the objects related to those resources.
- *Domain layer:* This layer deals with the ontologies used for for malizing the semantic information for both the DL pages (DC vocabulary) and contents (RDFS vocabulary or vocabularies).
- *Application layer:* This layer is supported on top of the previous one and will be built with the domain ontologies the Calli machusDL system requires, and applied to the resources in the first layer.

Keeping these knowledge layers conceptually separated, imple mentation will guarantee the flexibility and reusability of the Cal limachusDL application for every type of domain. Fig. 1 shows the framework of this approach, along with named examples for better understanding.

The domain layer holds the different domain ontologies that can be used. The application layer will use one or more domain ontologies depending on the sort of topics the DL application is going to deal with. The Dublin Core Ontology will always be used to represent the basic metadata concepts of every resource.

4.3. Using CallimachusDL

The CallimachusDL implementation is based on the SWAN architecture successfully deployed on CoolWikiNews, a Semanti cally enhanced Wiki devoted to online news publishing (Fuentes Lorenzo et al., 2007). CallimachusDL implements the Model View Controller (MVC) architectural pattern by means of Ruby on Rails (RoR, http://www.rubyonrails.org), a MVC based frame work which eases the task of building this architectural pattern. The common ontology used for describing the resources is Dublin Core. Its terms allow defining the metadata related to the whole page. The MarcOnt ontology is used for the annotation of more complex data. Both ontologies are developed with RDF Schema, and serialized in N Triple syntax (http://www.w3.org/TR/rdf test cases/#ntriples). The DL pages are presented to the user in XHTML 1.0 syntax (http://www.w3.org/TR/xhtml1/), and visual graphics for navigation are made with JavaScript libraries such as CoolTip (http://www.acooltip.com). Persistence repositories are MySQL server for resources information and SQLite based RDFLite for semantic information. Finally, CoolWikNews uses ActiveRDF (Oren, Delbru, & Decker, 2006), a library for abstracting the queries for RDFLite within the implementation in RoR.
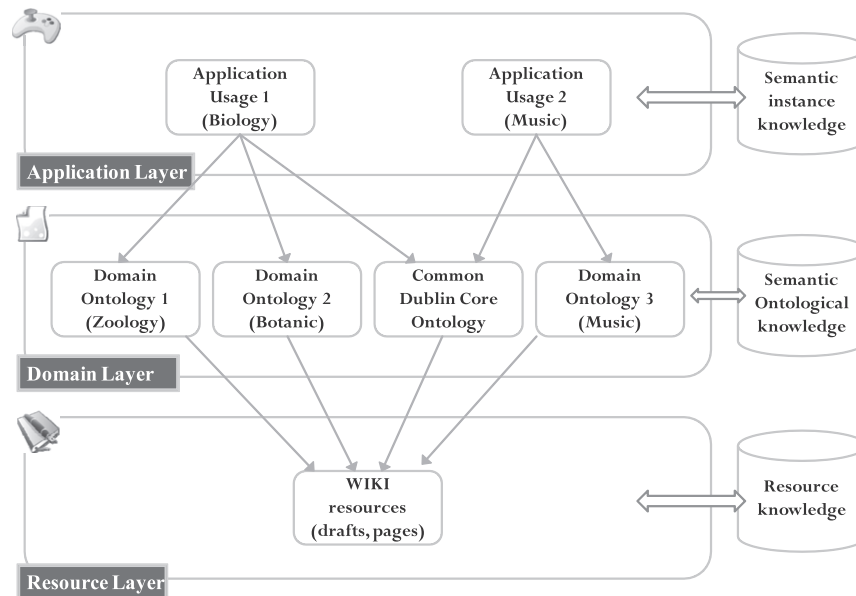
Fig. 1. The SWAN architecture as a basis for the CallimachusDL architecture.

A motivating scenario will now be presented to illustrate how CallimachusDL can be used. Recently, a new breed of user gener ated content aware technologies which have been encompassed by the "Web 2.0" buzzword have turned up to provide a huge amount of metadata and information about the user as a particular entity.

Web 2.0 technologies as outlined in (Laudon & Laudon, 2006) are exemplified by blogs, namely easy to update websites about a particular subject where entries are written in chronological or der, picture sharing environments such as Flickr or Photobucket, social bookmarking sites such as Del.icio.us, video sharing such as YouTube or music preferences such as Last FM. Web 2.0, social software, social computing, online communities, peer networking, immersive Web. Their meanings overlap, and definitions are some what fluid (Parameswaran & Whinston, 2007). But according to O'Reilly (2007) the term Web 2.0 is slightly different in that it in cludes more technologies within its scope and does not bind itself closely with the social aspect, as shown in several recent works (e.g. Correa, Willard Hinsley, & Gil de Zúñiga, 2010; Kwon & Wen, 2010). This Web 2.0 user generated content is a perfect sce nario to demonstrate the use of CallimachusDL, which also follows previous efforts in which multimedia contents are classified using semantics (e.g. Labra Gayo, Ordóñez de Pablos, & Cueva Lovelle, 2010).

For example, a user called John Smith has uploaded a number of videos in YouTube about his stay in Norway. In particular, the vid eos are about Norwegian Fjords so he tags them with the "fjord" and "Norway" tags. However, tags are freely chosen keywords describing a particular resource. They offer a simple way of retriev ing content but they are subjective conceptualizations, being potentially aggregated to a flat bottom up categorization or folks onomy. In Shadbolt, Hall, and Berners Lee (2006), folksonomies have been claimed to be an interesting emergent attempt for infor mation retrieval but serve different purposes to ontologies, the lat ter are attempts to more carefully define parts of the data world and to allow mappings and interactions between data held in dif ferent formats. In this scenario folskonomies had been used for cre ating semantic metadata (Al Khalifa & Davis, 2007) or as a support to learning (Lux & Dosinger, 2007). Hence, ontologies are defined through a careful, explicit process that attempts to remove ambi guity, whereas the definition of a tag is a loose and implicit process

where ambiguity might well remain. Finally, the inferential pro cess applied to ontologies is logic based and uses operations such as join. The inferential process used on tags is statistical in nature and employs techniques such as clustering (see Fig. 2).

If John Smith chooses any traditional DL, he will face a number of problems, as shown in Section 3. First of all, there is no metadata description, no chance of faceted browsing and problems to locate and retrieve his YouTube videos. Nevertheless, CallimachusDL of fers a completely different situation. Using the three layered archi tecture described in Section 4.2, the Resource Layer would store references to the videos in YouTube or the videos as such. In the Domain Layer, there are metadata formally describing the videos by means of ontologies, mostly DC and the MarcOnt ontology. After that, the Application Layer will use domain ontologies (for exam ple, those referring to Norway and Fjords).

Finally, faceted search and browsing would make the life of John Smith very easy when retrieving his videos, since he can nav igate through the categories and also view related videos thanks to the semalinks, as explained in Section 3.

5. Evaluation

The subsequent section describes the empirical evaluation of CallimachusDL, carried out using the application of established sta tistical tests.

5.1. Research design

Evaluation of the proposal was required in order to determine the level of acceptance of the proposal. With the objective of calcu lating the grade of adjustment of the proposal, a study was de signed which was aimed towards students in the final year of the Computer Science degree of the University Carlos III, specifically, those undertaking the subject "Software Engineering III". Among its learning objectives, this subject has the aim that the students are capable of extracting the requirements of a software applica tion of a client, with the final objective of being able to construct it according to a defined methodology. Therefore, the students were requested to use CallimachusDL to categorise distinct con tents generated by the students themselves during the course of
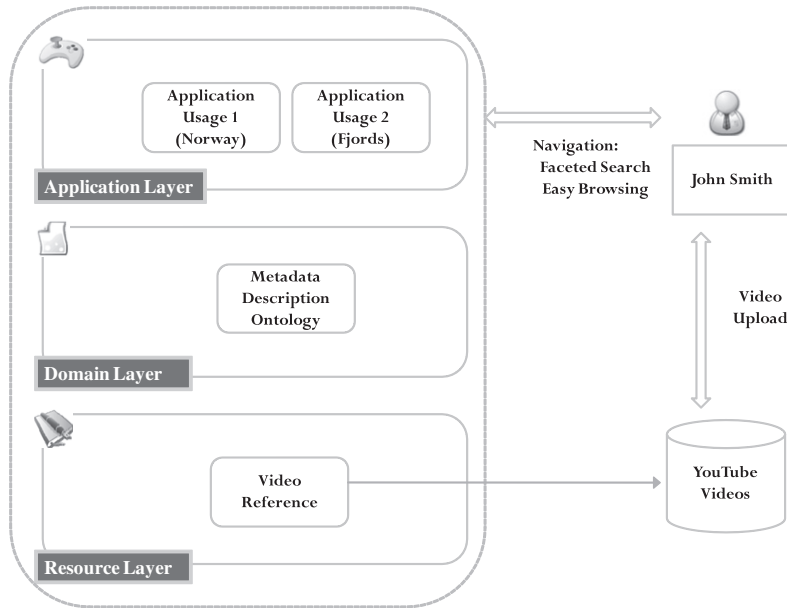
4

Fig. 2. John Smith videos in CallimachusDL.

completion of the subject. In the first place, the students carried out two interviews with the final clients (a role represented by one of the lecturers). The multimedia formats generated were up loaded to YouTube. In the second place, the students visited one of the work locations of the client, where they photographed the installations as part of the modeling of the current system. The photographs were uploaded to Flickr for sharing among the mem bers of the work team. Thirdly, the students were instructed to share the links which were Internet resources related to Software Engineering using the tool Del.icio.us. The use of these three Web 2.0 related technologies were the basis for the evaluation of CallimachusDL.

Once the students had completed the subject, they were re quested to respond to a questionnaire in which they were asked to provide their opinion about the tool in relation to different as pects. Firstly, they were asked about their user experience of Calli machus DL. Secondly, the user provided his opinion about the performance of the framework. The responses to these questions were coded using a Likert scale ranging from 1 to 4 points, with the following values. 1: Limited, 2: Regular, 3: Good, 4: Very Good. Lastly, it was suggested that each user carried out two types of search. The first search required the user to perform a keyword search, and the second search proposed to the user was to perform a search using the interface which implements faceted search. Upon examining the results of the search, the user could quantify the result using a Likert scale with the following values: 1: Un solved, 2: Low matching, 3: Near matching, 4: Correct matching.

5.2. Sample

The sample was composed of students in the final year of the Computer Science degree of the University Carlos III. The students, as part of their tasks within the subject "Software Engineering III", used CallimachusDL as support for the classification and annota tion of diverse contents, related to the themes of the subject.

The sample was composed of 17 women (32%) and 35 men (68%), with an average age of 25.6. During the administration of the questionnaire, the subjects were aided by a research assistant with the objective that the coding of the questionnaire was carried out accurately, eliminating any student doubts or errors.

5.3. Results

The results of the surveys, which were performed using printed copies, were subsequently coded in the statistical analysis tool SPSS. In the tables displayed below, the average and standard devi ation of the responses offered by the students are shown in relation to the questionnaire applied, and the two groups of questions formulated.

Fig. 3 shows the relative importance of the scores by factor and value in the Likert scale (see Table 1).

With the objective of verifying whether faceted search obtained results significantly higher than those of keyword search, the sta tistical method Student's t test (comparison of two means) was used to carry out one way between groups analysis of variance using the tool SPSS. The level of statistical significance was set at 0.05. The results of the test indicate that the Faceted Search tool presents significant differences with respect to Keyword Search indicated by the statistical value ($t(52) = 4.18$, $p < .05$). This circum stance implies that, from a statistical point of view, there is a dif ference between the faceted search and the keyword search. In this case, this is represented as an improvement in the results of faceted search with respect to keyword search.

Additionally, it was attempted to establish whether the vari ables "Faceted Search Experience" and "Faceted Search" presented
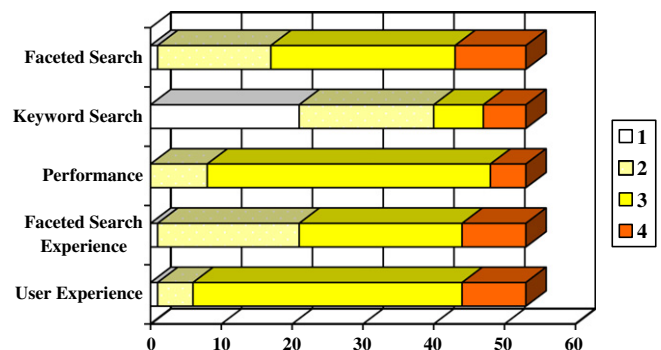

Fig. 3. Scores bar chart.

5

Table 1
Evaluation statistical results.

	Average	SD
User Experience	3.04	0.59
Faceted Search Experience	2.75	0.76
Performance	2.94	0.49
Keyword Search	1.96	1.00
Faceted Search	2.85	0.74

consistent results with respect to each other. "Faceted Search Experience", obtained the opinion of the user in relation to the perception of the feature, and "Faceted Search" included the search results. Similarly to the statistical analysis previously mentioned, the Student's t test was applied, setting the level of statistical significance at 0.05. By conventional criteria, the difference is considered to be not statistically significant ($t(52) = 0.65$, $p > .05$). This circumstance indicates that the opinion of the users in relation to faceted search is coherent with the opinion expressed regarding the results of the search.

5.4. Discussion

From the perspective of the results, the implementation of CallimachusDL may be considered a success. The opinions of the user in relation to "User Experience", "Faceted Search Experience" and "Performance" have displayed satisfactory results in average figures. In particular, the factor "User Experience" displays a score of 3.04 points over 4, with a highly adjusted standard deviation of 0.59, which indicates that, apart from the scoring being high, the agreement between the subjects is acceptable. The attempts to develop the Graphical User Interface of CallimachusDL so that it would be attractive to users has been judged positively by test users.

In relation to the searches carried out by the subjects, the judgments of the faceted search present higher values than the keyword search. Taking into account that the users have experience with keyword search, the higher ranking of faceted search in comparison with keyword search is an extremely significant result. Apart from the statistical evidence provided by the statistical test which evaluated the improvement, it can be indicated that the performance of faceted search, the differentiating element of CallimachusDL over other platforms, has been highly strengthened.

Lastly, it should be indicated that the results of the variables "Faceted Search Experience" and "Faceted Search" are very similar, not presenting significant differences between both. Evidently, the results are not identical, the second variable displaying a difference of 0.10 positive points with respect to the first. This circumstance can be explained by the differences of scales which both variables represent.

6. Conclusions and future work

Callimachus (c.305 c.240 B.C.) was an ancient Greek poet, librarian, and scholar, a famous representative of the Alexandrian school of poetry. Following the works of Zenodotus of Ephesus, Alexandria Library's first library director that began an inventory of the scrolls acquired by the Ptolemies, Callimachus created for the first time a subject catalog in 120,000 scrolls of the Library's holdings, called the Pinakes or Tables (Bevan, 1968). Following the Callimachus efforts, the man that improved subject search in Alexandria, this paper presents a novel approach to improve browsing and searching in DL by adding semantics to the definition of resources, particularly multimedia resources. In a larger context, the problem of DL scaling may be multiplied by thousands of data structures located in hundreds of incompatible databases and mes

sage formats. The uses of semantic DLs and social networks have been identified as one of the most promising research line in the DLs field (Kruck & McDaniel, 2008). In this scenario, CallimachusDL can be seen as an original initiative to bridge the gap between DLs and social web.

Hence, future work will consist of evaluating the implementation and approach more carefully, validating CallimachusDL with a number of quality aware case studies and using large DL resources where pooling out of results can determine more accurately if the effectiveness of the breakthroughs of the approach detailed in Section 3 take place. In a more general view, future work should further integrate social networks full potential into DLs. In particular, the increasing organizational capacities of distinct Web 2.0 tools represent a challenge for CallimachusDL, which in the future should be able to combine different multimedia contents generated from various platforms, to act as support for various tasks such as Customer Relationship Management (CRM), Talent Management (TM) or Knowledge Management (KM). As Galinec (2010) stated, supporting the acquisition and then deliver it to one or more target databases could improve the overall maturity of the IT function.

The unlimited potential of the Web 2.0 is an open field for technology investigators around the globe, and it is also a great opportunity for DLs researchers to put together social features and limitless content into a single package.

References

Al-Khalifa, H. S., & Davis, H. C. (2007). Exploring the value of folksonomies for creating semantic metadata. *International Journal on Semantic Web and Information Systems, 3*(1), 13–39.

Baruzzo, A., Casoto, P., Challapalli, P., Dattolo, A., Pudota, N., & Tasso, C. (2009). Toward Semantic Digital Libraries: Exploiting Web2.0 and Semantic Services in Cultural Heritage. *Journal of Digital Information, 10*(6). p. 4.

Bevan, E. (1968). *The house of ptolemy*. Chicago: Argonaut Inc.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American, 284*(5), 34–43.

Berners-Lee, T. (2007). Giant Global Graph. Available from http://dig.csail.mit.edu/breadcrumbs/node/215, November 2007. Last Verified: January 15th, 2010.

Bygstad, B., Ghinea, G., & Klæboe, G. T. (2009). Organisational challenges of the semantic web in digital libraries: A Norwegian case study. *Online Information Review, 33*(5), 973–985.

Colomo-Palacios, R., Gómez-Berbís, J. M., García-Crespo, A., & Puebla Sánchez, I. (2008). Social Global Repository: using semantics and social web in software projects. *International Journal of Knowledge and Learning, 4*(5), 452–464.

Correa, T., Willard Hinsley, A., & Gil de Zúñiga, H. (2010). Who interacts on the Web?: The intersection of users' personality and social media use. *Computers in Human Behavior, 26*(2), 247–253.

Davies, J., Lytras, M. D., & Sheth, A. P. (2007). Semantic-web-based knowledge management. *IEEE Internet Computing, 11*(5), 14–16.

Fensel, D. (2002). *Ontologies: A silver bullet for knowledge management and electronic commerce*. Berlin: Springer.

Ferrán, N., Mor, E., & Minguillón, J. (2005). Towards personalization in digital libraries through ontologies. *Library Management, 26*(4/5), 206–217.

Fuentes-Lorenzo, D., Gómez-Berbís, J. M., & García-Crespo, Á. (2007). CoolWikNews: More than meets the eye in the XXI century journalism. In J. Rech, B. Decker & E. Ras (Eds.), *Emerging technologies for semantic work environments: Techniques, methods, and applications*, IGI-Global.

Fuentes-Lorenzo, D., Morato, J., & Gómez-Berbís, J. M. (2009). Knowledge management in biomedical libraries: A semantic web approach. *Information Systems Frontiers, 11*(4), 471–480.

Galinec, D. (2010). Human capital management process based on information technology models and governance. *International Journal of Human Capital and Information Technology Professionals, 1*(1), 44–60.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J. M., & García-Sánchez, F. (2010). SOLAR: Social link advanced recommendation system. *Future Generation Computer Systems, 26*(3), 374–380.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J. M., & Mencke, M. (2009). BMR: Benchmarking Metrics Recommender for Personnel issues in Software Development Projects. *International Journal of Computational Intelligence Systems, 2*(3), 257–267.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J. M., & Ruiz-Mezcua, B. (2010). SEMO: A framework for customer social networks analysis based on semantics. *Journal of Information Technology, 25*(2), 178–188.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220.

Han, Y. (2006). A RDF-based digital library system. *Library Hi Tech, 24*(2), 234–240.

Hendler, J. (2008). Web 3.0: Chicken Farms on the Semantic Web. *Computer, 41*(1), 106–108.

Isaac, A., Schlobach, S., Matthezing, H., & Zinn, K. (2008). Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies. *Library Review, 57*(3), 187–199.

Kim, H., & Choi, C. (2000). XML how it will be applied to digital library systems. *The Electronic Library, 18*(3), 183–189.

Kruk, S.R., & Decker, S. (2007). JeromeDL: The Semantic Digital Library. In *Proceedings Semantic Technology Conference 2007*, San José, California.

Kruk, S.R., Decker, S., & Zieborak, L. (2005). JeromeDL: Adding Semantic Web Technologies to Digital Libraries. In *Proceedings of the 16th International conference on database and expert systems applications*, Copenhagen, Denmark.

Kruck, S. R., & McDaniel, B. (2008). *Semantic Digital Libraries*. Berlin Heidelberg: Springer.

Kruk, S. R., Synak, M., & Zimmermann, K. (2005). MarcOnt: Integration ontology for bibliographic description formats. In *Proceedings of the International Conference on Dublin Core and Metadata Applications (DC-2005)*, Madrid, Spain, September 2005.

Kwon, O., & Wen, Y. (2010). An empirical study of the factors affecting social network service use. *Computers in Human Behavior, 26*(2), 254–263.

Labra Gayo, J. E., Ordóñez de Pablos, P., & Cueva Lovelle, J. M. (2010). WESONet: Applying semantic web technologies and collaborative tagging to multimedia web information systems. *Computers in Human Behavior, 26*(2), 205–209.

Lassila, O., & Hendler, J. (2007). Embracing "Web 3.0". *IEEE Internet Computing, 11*(3), 90–93.

Laudon, K. C., & Laudon, J. P. (2006). *Management Information Systems: Managing the Digital Firm* (10th ed.). Upper Saddle River, NJ: Prentice Hall.

Lux, M., & Dosinger, G. (2007). From folksonomies to ontologies: Employing wisdom of the crowds to serve learning purposes. *International Journal of Knowledge and Learning, 3*(4/5), 515–528.

Lytras, M. D., & García, R. (2008). Semantic Web applications: a framework for industry and business exploitation – What is needed for the adoption of the Semantic Web from the market and industry. *International Journal of Knowledge and Learning, 4*(1), 93–108.

Lytras, M. D., Damiani, E., & Ordóñez de Pablos, P. (2008). *Web 2.0: The Business Model*. Springer.

Lytras, M., & Ordóñez de Pablos, P. (2007). Red Gate Corner: a Web 2.0 prototype for knowledge and learning concerning China business and culture. *International Journal of Knowledge and Learning, 3*(4&5), 542–548.

Lytras, M. D., & Ordóñez de Pablos, P. (2009). Social Web Evolution. Integrating Semantic Applications and Web 2.0 Technologies. IGI-Global.

Mazurek, C., & Werla, M. (2005). Distributed services architecture in dLibra digital library framework. In: 8th International Workshop of the DELOS Network of Excellence on Digital Libraries on Future Digital Library Management Systems.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh: Edinburgh University Press.

O'Reilly, T. (2007). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Communications & Strategies, 65*, 17–27.

Ordóñez de Pablos, P. (2002). Knowledge management and organizational learning: Typologies of generic knowledge strategies in the Spanish manufacturing industry from 1995 to 1999. *Journal of Knowledge Management, 6*(1), 52–62.

Oren, E., Delbru, R., & Decker, S. (2006). Extending faceted navigation for RDF data. In *Proceedings of the International Semantic Web Conference, Lecture Notes in Computer Science* (pp. 559–572). Berlin: Springer.

Parameswaran, M., & Whinston, A. B. (2007). Research issues in social computing. *Journal of the Association of Information Systems, 8*(6), 336–350.

Prasad, A. R. D., & Madalli, D. P. (2008). Faceted infrastructure for semantic digital libraries. *Library Review, 57*(3), 225–234.

Ranganathan, S. R. (1962). *Elements of library classification*. New York: Asia Publishing House.

Rodriguez Pérez, J. M., & Ordóñez de Pablos, P. (2003). Knowledge management and organizational competitiveness: A framework for human capital analysis. *Journal of Knowledge Management, 7*(3), 82–91.

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The Semantic Web Revisited. *IEEE Intelligent Systems, 21*(3), 96–101.

Schreiber, G., Amin, A., Aroyo, L., van Assem, M., de Boer, V., Hardman, L., et al. (2008). Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator. *Web Semantics: Science, Services and Agents on the World Wide Web, 6*(4), 243–249.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *IEEE Transactions on Data and Knowledge Engineering, 25*(1–2), 161–197.

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: A study of orienteering behaviour in directed search. In: *Proceedings of the ACM conference on human factors in computing systems*, Vienna, Austria, pp. 415–422.

Vossen, G., Lytras, M., & Koudas, N. (2007). Editorial: Revisiting the (Machine) Semantic Web: The Missing Layers for the Human Semantic Web. *IEEE Transactions on Knowledge and Data Engineering, 19*(2), 145–148.

Warren, P. (2006). Knowledge management and the semantic web: From scenario to technology. *IEEE Intelligent Systems, 21*(1), 53–59.