Krassimir Markov, Vitalii Velychko,
Lius Fernando de Mingo Lopez, Juan Casellanos
(editors)

New Trends in Information Technologies

I T H E A
SOFIA
2010

Krassimir Markov, Vitalii Velychko, Lius Fernando de Mingo Lopez, Juan Casellanos (ed.)

New Trends in Information Technologies

ITHEA®

Sofia, Bulgaria, 2010

ISBN 978-954-16-0044-9

First edition

Recommended for publication by The Scientific Concil of the Institute of Information Theories and Applications FOI ITHEA

This book maintains articles on actual problems of research and application of information technologies, especially the new approaches, models, algorithms and methods of membrane computing and transition P systems; decision support systems; discrete mathematics; problems of the interdisciplinary knowledge domain including informatics, computer science, control theory, and IT applications; information security; disaster risk assessment, based on heterogeneous information (from satellites and in-situ data, and modelling data); timely and reliable detection, estimation, and forecast of risk factors and, on this basis, on timely elimination of the causes of abnormal situations before failures and other undesirable consequences occur; models of mind, cognizers; computer virtual reality; virtual laboratories for computer-aided design; open social info-educational platforms; multimedia digital libraries and digital collections representing the European cultural and historical heritage; recognition of the similarities in architectures and power profiles of different types of arrays, adaptation of methods developed for one on others and component sharing when several arrays are embedded in the same system and mutually operated.

It is represented that book articles will be interesting for experts in the field of information technologies as well as for practical users.

General Sponsor: Consortium FOI Bulgaria (www.foibg.com).

Printed in Bulgaria

ISBN 978-954-16-0044-9

C\o Jusautor, Sofia, 2010

SEARCH AND ADMINISTRATIVE SERVICES IN ICONOGRAPHICAL DIGITAL LIBRARY

Desislava Paneva-Marinova, Radoslav Pavlov, Maxim Goynov, Lilia Pavlova-Draganova, Lubomil Draganov

***Abstract**: Today there are a large number of digital archives, libraries and museums with rich digital collections representing the European cultural and historical heritage. The new challenge shifts from having online access to resources to making an effective use of them and avoiding information overload. A possible answer to this challenge is the approach, used for the development of the "Virtual encyclopedia of the Bulgarian Iconography" digital library (BIDL). It is a complete web-based environment for registration, documentation, access and exploration of a practically unlimited number of Bulgarian iconographical artefacts and knowledge. The key for its efficiency is the provision of strictly designed functionalities, powered by a long-term observation of the users' preferences, cognitive goals, and needs, aiming to find an optimal functionality solution for the end users. A special attention was pay to search and administrative services, trying to cover a wide range of possible solutions such as keyword search, extended keyword search, semantic-based search, complex search, search with result grouping, tracking services, DL data exportation, etc. This paper presents these services in detail, their functional specifications and used algorithms. The ontology of the East-Christian Iconographical Art is discussed, because of its important role for the semantic description and search of iconographical artefacts and knowledge in the library.*

***Keywords**: multimedia digital libraries, systems issues, user issues, online information services*

***ACM Classification Keywords**: H.3.5 Online Information Services – Web-based services, H.3.7 Digital Libraries – Collection, Dissemination, System issues.*

Introduction

In an attempt to answer the need for presentation and preservation of the Bulgarian iconography, a team from the Institute of Mathematics and Informatics has developed a multimedia digital library called Virtual Encyclopedia of Bulgarian Iconography (http://mdl.cc.bas.bg). It was designed so as to provide wide accessibility and popularization of the works of the Bulgarian iconographers, and moreover to enable future precise restoration of the icons at risk.

The "Virtual encyclopedia of the Bulgarian Iconography" digital library (also called Bulgarian Iconography Digital Library, BIDL)[1] is a complete web-based environment for registration, documentation, access and exploration of a practically unlimited number of Bulgarian iconographical artefacts and knowledge. It provides a rich knowledge base for the iconographical art domain, enabling its usage for content annotation, preview, complex search, selection, group and management. The ontology of the East-Christian iconographical art was developed and used for semantic annotation of the library content [Pavlov et al., 2010] [Pavlova-Draganova et al., 2007b] [Paneva et al., 2007].

[1] The first release of the BIDL was developed five years ago during the project "Digital Libraries with Multimedia Content and its Application in Bulgarian Cultural Heritage" (contract 8/21.07.2005 between the Institute of Mathematics and Informatics, BAS, and the State Agency for Information Technologies and Communications), aiming to lay the foundations of the registration, documentation, and the exploration of a practically unlimited number of Bulgarian icons [Pavlova-Draganova et al., 2007a] [Pavlov et al., 2006].

A very important task during the BIDL development was the provision of the strictly designed functionalities. A special attention was pay to search and administrative services, trying to cover a wide range of possible solutions such as keyword search, extended keyword search, semantic-based search, complex search, search with result grouping, tracking services, DL data exportation, *etc.*

This paper extends the BIDL functionality presentation of [Pavlov et al., 2010], where the content creation and preview services was mainly discussed. The paper structure is the following: Section 2 makes a short overview of the BIDL architecture, covering its main service panels, repositories and their relationships. The search and administrative services are presented in details in Section 4 and Section 5, where their functionality and base algorithms are described. The ontology of the East-Christian Iconographical Art knowledge is discussed in Section 3, because of its key role for the semantic search of iconographical artefacts and knowledge in the library. Section 5 summarizes the achieved results and traces the directions for future development of BIDL.

BIDL Architecture

The BIDL environment, depicted in figure 1, integrates:

- Appropriate repositories and services for management of two types of objects:
    - MDL objects – multimedia digital objects, described by technical and semantic metadata and saved in a *Media Repository*;
    - User profiles, presenting user's data and behavior, saved in a User Profile Repository.
- Two main service panels, named *Object data management* and *Administrative services*, provide the base BIDL functionality.

The *Object data management* panel refers to the activities related to: content creation: add (annotate and semantic indexing), store, edit, preview, delete, group, and manage multimedia digital objects; manage metadata (see [Pavlov et al., 2010]); search, select (filter), access and browse digital objects, collections and their descriptions. The *Administrative services* panel mainly provides user data management, data export, tracing and analysis services, presented bellow.

For every MDL object all semantic and technical metadata are saved in the Media repository. These metadata are represented in catalogue records that point to the original media file/s associated to every MDL object. The User profile repository manages all user data and their changes.

There are several internal relations between the separate components in the service panels. For example, in the Object data management panel:

- the *Add object services* are related to the *Preview* and *Edit services*;
- after the *Preview (services)*, the *Edit* or *Delete services* can be executed;
- the Search object services point to *Preview, Edit, Delete* and *Group objects services*;
- the *Group objects services* are related to *Preview services*;
- after the *Edit (services)*, *the Preview services* can be executed.

There are several relations between the components of the two main service panels, for example, the Tracing of MDL objects from the Administrative services panel is connected to Add object, Preview, Delete, Search, Edit and Group services from the Object data management panel.

All existing internal and external relations for the service panels provide the internal interoperability and the flexibility of the library.

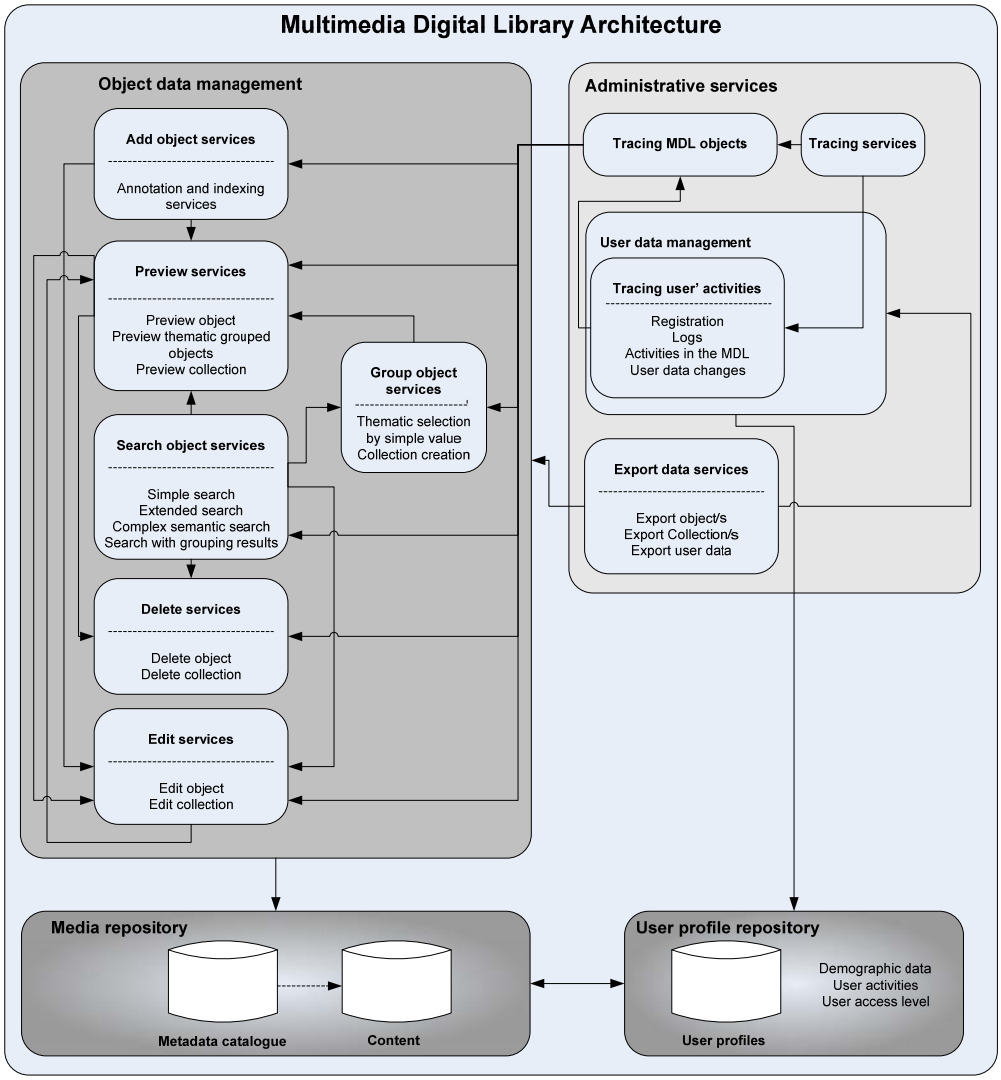

Figure 1: BIDL Architecture

Semantic Description of Iconography Art World

The semantic metadata description of Bulgarian icon art is determined by the domain ontology of the East-Christian iconographical art (also called iconography ontology)[1]. It presented the iconographical art world by three "thematic entities"[2] (also called levels of knowledge). Every one of these entities is enriched with a set of sub-levels, covering wide range of characteristics. The first one is the "Identification" entity (see Figure 2), which consists of general data identifying aspects such as IO title, type, author and biographical data for the object's author, its clan, iconographic school, period, dimensions, current location and source, and object identification notes, iconographic school description.

[1] The ontology of the East-Christian iconographical art was developed for resource semantic annotation for the project SINUS "Semantic Technologies for Web Services and Technology Enhanced Learning" (№ D-002-189).

[2] A development methodology used during the creation of the *Ontology of the iconographical objects (artefacts)* [Paneva et al., 2007] [Pavlova-Draganova et al., 2007b].

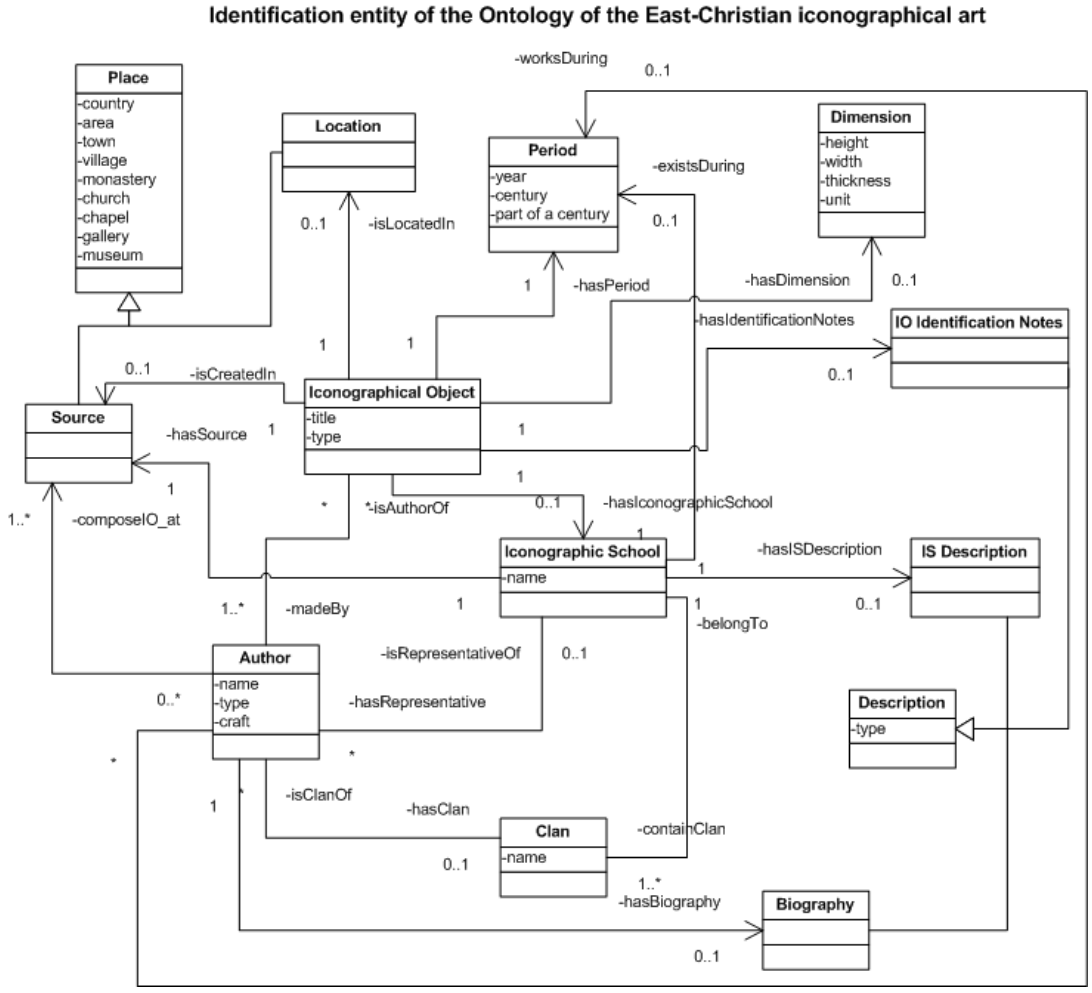

Figure 2: Identification entity of the Ontology of the East-Christian iconographical art

The second entity (see Figure 3) covers information concerning the descriptive details of the theme and forms of representation, providing a better understanding of the content.

The main concepts included are: depicted character/s, iconographical scenes, character/s in the scene/s, symbol/s in the scene/s, character' gestures, character' vestment, detailed description of the depicted content, *etc.* The third entity covers technical information revealing iconographic techniques, base materials, varnishes, gilding, *etc.*, used in the creation of the iconographical object/collection, and also concerning examinations of the condition, such as diagnosis or history of the conservation treatment.

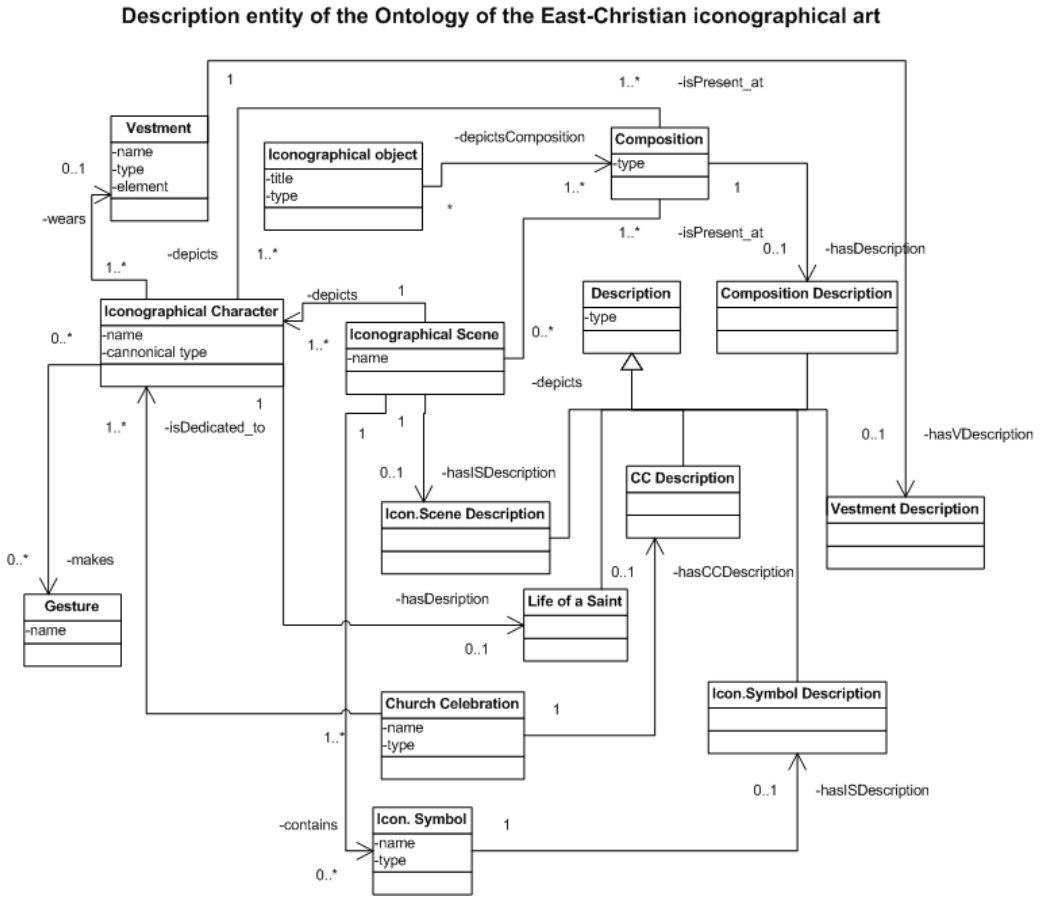

Figure 3: Description entity of the Ontology of the East-Christian iconographical art

Search Services

BIDL provides a wide range of search services, such as keyword search, extended keyword search, semantic-based search, complex search, and search with grouping results. This section presents the complex search algorithm that is base of all other search possibilities.

Let $U = O \times C$, O is the set of objects and C is the set of characteristics and U is the set of all objects and their characteristics. Let $v(o,c)$ is a function : $v : O \times C \to V$, where V is the set of values of the characteristics.

$p(c,v)$ is a condition for the characteristic c and the value v. In the first version of our search service, there was only one type of condition: $p(c,v) \Leftrightarrow$ "objects having value v for characteristic c". Let P be the set of all possible conditions for $c \in C$ and $v \in V$.

Let define the search function $s(p,u)$, where $p \in P$ and $u \in U$, $s : P \times U \to U$. The result is a set $S \subseteq U$.

Let assume that we search on n characteristics.

So, in the first version of our searching service, we used the following algorithm:

$S_1 = s(p_1, U) \rightarrow$ time for execution $= t_1 = t$

$S_2 = s(p_2, U) \rightarrow$ time for execution $= t_2 \approx t$

$S_3 = s(p_3, U) \rightarrow$ time for execution $= t_3 \approx t$

. . .

$S_n = s(p_n, U) \rightarrow$ time for execution $= t_n \approx t$,

where p_n are the conditions for all n characteristic.

The result of our search will be:

$R = S_1 \cap S_2 \cap S_3 \ldots \cap S_n$ (See figure 4)

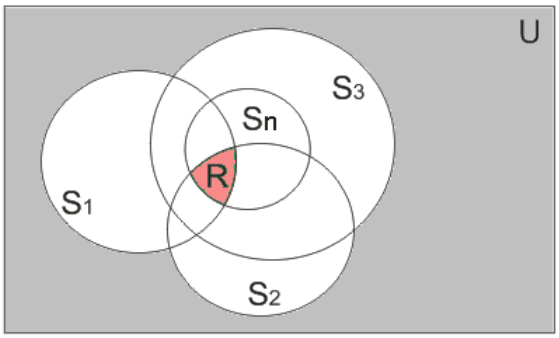

Figure 4: Result set of the search in the first release of the BIDL

If we assume that the time for making one search iteration over the all set of objects and their characteristics U is t, therefore the execution of the whole algorithm will spend $t_{v1} = t_n$ time + the time needed for the intersection of the results in the first release of searching services.

The current version of the searching service had the following changes:

The types of conditions raised to 5:

1. "objects having value $= v$ for characteristic c" – the same as in version 1
2. "objects having value $\neq v$ for characteristic c"
3. "objects having numeric value $\geq, \leq, <, >, or = v$ for characteristic c"
4. "objects having characteristic c"
5. "objects NOT having characteristic c"

The algorithm for the search function changed to (see figure 5):

$S_1 = s(p_1, U) \rightarrow$ time for execution $= t_1 = t$

$S_2 = s(p_2, S_1) \rightarrow$ time for execution $= t_2 \leq t_1$

$S_3 = s(p_3, S_2) \rightarrow$ time for execution $= t_3 \leq t_2$

. . .

$S_n = s(p_n, S_{n-1}) \rightarrow$ time for execution $= t_n \leq t_{n-1}$

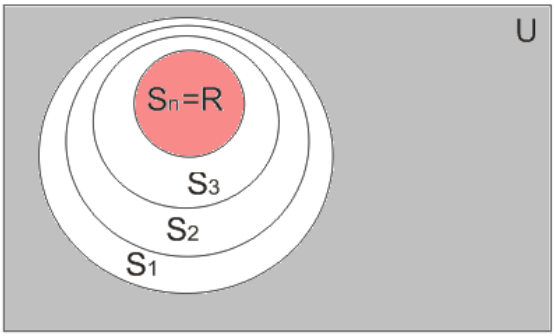

Figure 5: Set of results of the search in the current release of the BIDL

The result R will be equal to S_n, so no intersection will be needed. The time for execution $t_i \leq t_{i-1}$ is because at each iteration the search set $S_i \subseteq S_{i-1}$, therefore the time for processing a search decreases.

In this way, the overall time for execution will be $t_{v2} = \sum_{i=1}^{n} t_i \Rightarrow t_{v2} \leq t_n$ and $t_{v2} < t_{v1}$, t_{v2} is the time needed for results generation in the current release of searching services.

Administrative Services

The *Administrative services* panel mainly provides user data management, data export, tracking services, and analysis services. The user data management covers the activities related to registration, data changes, level set, and tracking activities of the user. The tracking services have two main branches: tracking of MDL objects, tracking of MDL user' activities (example, figure 6). The tracking of MDL objects spies on the activities of add, edit, preview, search, delete, selection, export to XML, and group of MDL objects/collections in order to provide a wide range of statistic data (for frequency of service usage, failed requests, *etc.*) for internal usage and generation of inferences about the stable work (stability), the flexibility, and the reliability of the environment. The tracking of MDL user' activities spies user logs, personal data changes, access level changes and user behavior in the BIDL.

The QlickTech® QlinView® Business Intelligence[1] software is the analysis services provider. It is connected to the BIDL tracking services and objects data base by preliminary created data warehouse[1].

[1] Business Intelligence is an architecture and a collection of integrated operational as well as decision-support applications and databases that provide easy access to great amount of (business) data.

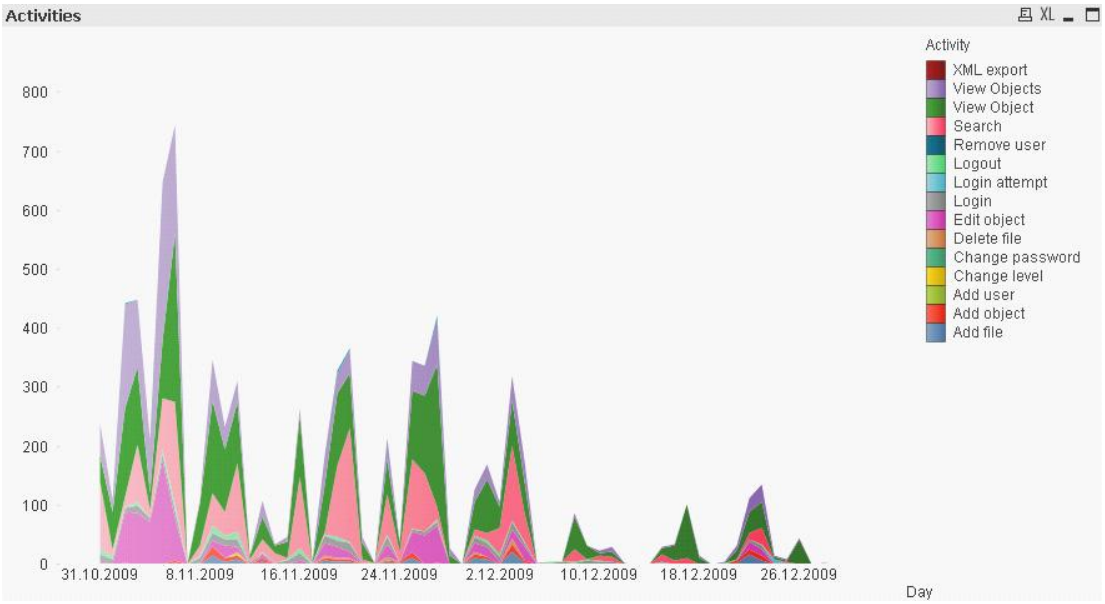

Figure 6: Users' activities during the period 10 – 12. 2009

The ETL (Extract, Transform, Load)[2] is completely automatic process and is performed by administrator request.

The QlickTech® QlinView® Business Intelligence Software is deployed in order to provide fast, powerful and visual in-memory analysis of the data in the warehouse. It is a data access solution that enables you to analyze and use information from different data sources. It is based on online analytical processing (OLAP), which provides an approach to quickly answer multi-dimensional analytical queries [Codd et al., 1993].

Figure 7 depicts PIE diagram making canonical sub-types analysis.

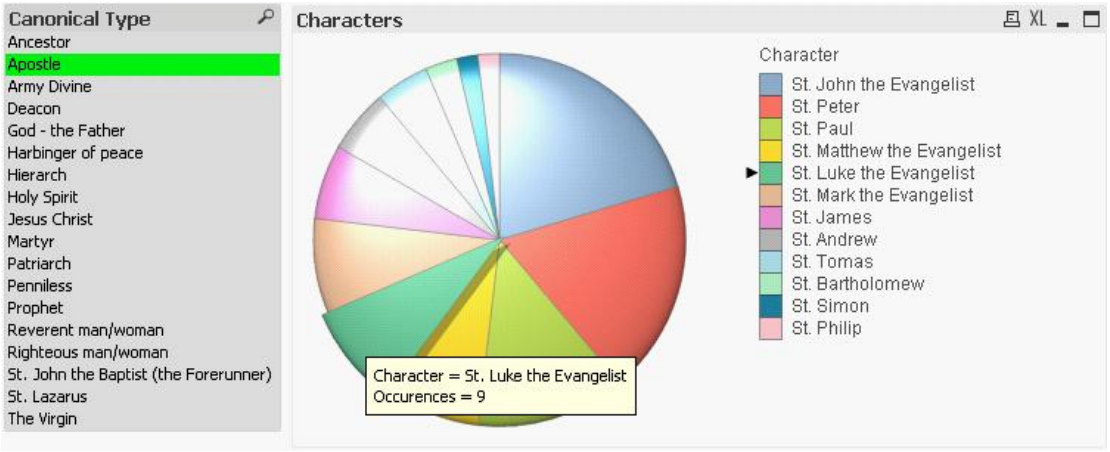

Figure 7: PIE diagram of canonical sub-types for Apostle canonical type

[1] A data warehouse is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and analysis [Inmon, 1995].

[2] Extract, transform, and load (ETL) is a process in database usage and especially in data warehousing that involves: extracting data from outside sources, transforming it to fit operational needs (which can include quality levels), and loading it into the end target (database or data warehouse).

The variety of generated statistic information about BIDL data using QlickTech® QlinView® provides a rich extension of the tracking services and the base for profound analysis of extracted data.

The export data from the administrative services panel provides the transfer of information packages (for example, packages with BIDL objects/collections, user profiles, *etc.*) compatible with other systems managing data bases. For example, with these services a package with BIDL objects could be transported in a XML-based structure for a new external usage in e-learning [Paneva-Marinova et al., 2008] [Pavlov and Paneva, 2007] [Paneva-Marinova et al., 2009] or e-commerce applications.

The following code shows an instance of object data, exported in XML.

```
<object id="1">
	<characteristics>
	<chr name="Идентификация" id="1_0_0" value="">
		<chr name="Заглавие" id="2_0_0" value="">
			<chr name="lang:bg" id="3_0_0" value="Св. Богородица Катафиги (убежище) и св. Йоан Богослов"/>
			<chr name="lang:en" id="4_0_0" value="The Virgin Cataphyge (Refuge) and St. John the Evangelist"/>
		</chr>
		<chr name="Тип на иконографския обект" id="5_0_0" value="">
			<chr name="Икона" id="6_0_0" value=""/>
		</chr>
		<chr name="Автор" id="20_0_0" value="">
			<chr name="lang:bg" id="21_0_0" value="неизвестен"/>
			<chr name="lang:en" id="22_0_0" value="Unknown"/>
		</chr>
		<chr name="Иконописна школа" id="29_0_0" value="">
			<chr name="lang:bg" id="30_0_0" value="неизвестна"/>
			<chr name="lang:en" id="31_0_0" value="Unknown"/>
		</chr>
		<chr name="Период" id="32_0_0" value="">
			<chr name="От" id="33_0_0" value="">
				<chr name="Година" id="34_0_0" value="1395"/>
			</chr>
		</chr>
		<chr name="Размери (см)" id="53_0_0" value="">
			<chr name="височина" id="54_0_0" value="93"/>
			<chr name="ширина" id="55_0_0" value="61.5"/>
		</chr>
		<chr name="Местонахождение" id="57_0_0" value="">
			<chr name="Държава" id="58_0_0" value="">
				<chr name="lang:bg" id="59_0_0" value="България"/>
				<chr name="lang:en" id="60_0_0" value="Bulgaria"/>
			</chr>
			<chr name="Област" id="61_0_0" value="">
				<chr name="lang:bg" id="62_0_0" value="София"/>
				<chr name="lang:en" id="63_0_0" value="Sofia"/>
			</chr>
			<chr name="Галерия" id="85_0_0" value="">
				<chr name="lang:bg" id="86_0_0" value="Национална художествена галерия"/>
				<chr name="lang:en" id="87_0_0" value="National Art Gallery"/>
			</chr>
		</chr>
	</chr>
	<chr name="Описание" id="128_0_0" value="">
		<chr name="Персонажи" id="129_0_0" value="">
			<chr name="Име на персонаж" id="130_0_0" value="">
				<chr name="lang:bg" id="131_0_0" value="Св. Богородица Катафиги (Убежище)"/>
				<chr name="lang:en" id="132_0_0" value="The Virgin Cataphyge (Refuge)"/>
			</chr>
			<chr name="Каноничен тип на персонаж" id="133_0_0" value="">
				<chr name="lang:bg" id="134_0_0" value="Св. Богородица"/>
				<chr name="lang:en" id="135_0_0" value="The Virgin "/>
			</chr>
		</chr>
	</chr>
	<chr name="Технология" id="146_0_0" value="">
		<chr name="Иконографска техника" id="147_0_0" value="">
			<chr name="lang:bg" id="148_0_0" value="Темперна"/>
			<chr name="lang:en" id="149_0_0" value="Tempera"/>
		</chr>
		<chr name="Основа" id="153_0_0" value="">
			<chr name="lang:bg" id="154_0_0" value="Дърво"/>
			<chr name="lang:en" id="155_0_0" value="Wood"/>
		</chr>
	</chr>
	</characteristics>
	<files>
		<file id="1" original_name="1.jpg" savedas="1.jpg" />
	</files>
</object>
```

Conclusions and Future Work

A tendency from the last few years points towards the use of digital libraries as a source of digital knowledge and environment for its delivery. This tendency determines the development of new methods and techniques for functionality provision, aiming to satisfy user's needs and preferences. The new MDL systems aim to find optimal functionality solutions, mainly by improving the content annotation, search and presentation, metadata management, environment administration, *etc.* In this paper we presented the core and the motor of the BIDL architecture: the search and administrative services, covering the used techniques and algorithms. The next step will be the implementation of personalized work space and dynamic content adaptation service, assisting individual user's content observation. A profound research is done for the provision of innovative techniques and tools for digital preservation and restoration of valuable artefacts of the Iconographical art world. The investigations are also directed towards the development of tools for aggregating iconographical content and ensuring its semantic compatibility with the European digital library EUROPEANA, thus providing possibilities for pan-European access to rich digitalised collections of Bulgarian Iconographical heritage.

Acknowledgements

This work is partly funded by Bulgarian NSF under the project D-002-189 SINUS "Semantic Technologies for Web Services and Technology Enhanced Learning".

This work is partially supported by Project BG051PO001/07/3.3-02/7 as a part of the grant scheme "Support for the Development of PhD Students, Post-doctoral Students, Post-graduate Students and Young Scientists" under the Operational programme "Human Resource Development" of the European Social Fund and the Bulgarian Ministry of Education and Science.

Bibliography

[Pavlov et al., 2010] Pavlov, P., D. Paneva-Marinova, M. Goynov, L. Pavlova-Draganova. Services for Content Creation and Presentation in an Iconographical Digital Library, International Journal "Serdica Journal of Computing", 2010 (in print)

[Paneva et al., 2007] Paneva, D., L. Pavlova-Draganova, L. Draganov. Towards Content-sensitive Access to the Artefacts of the Bulgarian Iconography, In the Proceedings of the Fifth International Conference "Information Research and Applications" – i.Tech 2007 (ITA 2007 - Xth Joint International Scientific Events on Informatics), 26 June – 01 July, 2007, Varna, Bulgaria, vol. 1, pp. 33-38.

[Paneva-Marinova et al., 2008] Paneva-Marinova, D., L. Pavlova-Draganova, R. Pavlov, M. Sendova. Cross-media and Ubiquitous Learning Applications on Top of Iconographic Digital Library. In the Proceedings of the 14th International Conference on Virtual Systems and Multimedia, Limassol, Cyprus, 20-25 October 2008, pp. 367-371.

[Paneva-Marinova et al., 2009] Paneva-Marinova D., L. Pavlova-Draganova, L. Draganov, R. Pavlov, M. Sendova. Development of a Courseware on Bulgarian Iconography for Ubiquitous On-demand Study. In: Szucs A. (Ed.) Proceedings of Open Conference "New Technology Platforms for Learning – Revisited". Budapest, Hungary, January 2009, pp. 37-46.

[Pavlov and Paneva, 2007] Pavlov R., D. Paneva. Toward Ubiquitous Learning Application of Digital Libraries with Multimedia Content, International Journal "Cybernetics and Information Technologies", vol. 6 (2007), № 3, pp. 51-62.

[Pavlova-Draganova et al., 2007a] Pavlova-Draganova L., V. Georgiev, L. Draganov. Virtual Encyclopaedia of Bulgarian Iconography, Information Technologies and Knowledge, vol.1 (2007), №3, pp. 267-271.

[Pavlova-Draganova et al., 2007b] Pavlova-Draganova, L., D. Paneva, L. Draganov. Knowledge Technologies for Description of the Semantics of the Bulgarian Iconographical Artefacts, In the Proceedings of the Open Workshop "Knowledge Technologies and Applications", Kosice, Slovakia, 31 May - 1 June, 2007, pp. 41-46.

[Pavlov et al., 2006] Pavlov R., L. Pavlova-Draganova, L. Draganov, D. Paneva, e-Presentation of East-Christian Icon Art, In the Proceedings of the Open Workshop "Semantic Web and Knowledge Technologies Applications", Varna, Bulgaria, 12 September, 2006, pp. 42-48.

[Pavlova-Draganova et al., 2009] Pavlova-Draganova, L. D. Paneva-Marinova. A Use Case Scenario for Technology-Enhanced Learning through Semantic Web Services, International Journal „Information Technologies & Knowledge", vol. 3, 2009 (in print)

[Codd et al., 1993] Codd E.F., S.B. Codd, C.T. Salley. Providing OLAP (On-line Analytical Processing) to User-Analysts: An IT Mandate. Codd & Date, Inc., 1993, pp. 1-31.

[Inmon, 1995] Inmon, W.H. Tech Topic: What is a Data Warehouse? Prism Solutions. Volume 1. 1995.

Authors' Information

Desislava Paneva-Marinova *– PhD in Informatics, Assistant Professor, Institute of Mathematics and Informatics, BAS, Acad. G. Bonchev Str.., bl. 8, Sofia 1113, Bulgaria; e-mail: dessi@cc.bas.bg*

Major Fields of Scientific Research: Multimedia Digital Libraries, Personalization and Content Adaptivity, eLearning Systems and Standards, Knowledge Technologies and Applications.

Radoslav Pavlov – *PhD in Mathematics, Associated Professor, Institute of Mathematics and Informatics, BAS, Acad. G. Bonchev Str.., bl. 8, Sofia 1113, Bulgaria; e-mail: radko@cc.bas.bg*

Major Fields of Scientific Research: Multimedia and Language Technologies, Digital Libraries, Information Society Technologies, e-Learning, Theoretical Computer Science, Computational Linguistics, Algorithmic, Artificial Intelligence and Knowledge Technologies.

Maxim Goynov – *Programmer, Institute of Mathematics and Informatics, BAS, Acad. G. Bonchev Str.., bl. 8, Sofia 1113, Bulgaria; e-mail: maxfm@abv.bg*

Major Fields of Scientific Research: Multimedia Digital Libraries and Applications.

Lilia Pavlova-Draganova *– Assistant Professor, PhD Student, Laboratory of Telematics, BAS, Acad. G. Bonchev Str., bl. 8, Sofia-1113, Bulgaria; e-mail: pavlova.lilia@gmail.com*

Major Fields of Scientific Research: Multimedia Digital Libraries, East Christian Iconography, 2D and 3D Computers Graphics and Animation, Design and Modelling Cultural Heritage Objects, Preservation and Restoration of Digitalized Artifacts, eLearning.

Lubomil Draganov *– Assistant Professor, PhD Student, Institute of Mathematics and Informatics, BAS, Acad. G. Bonchev Str.., bl. 8, Sofia 1113, Bulgaria; e-mail: lubomil@gmail.com*

Major Fields of Scientific Research: Multimedia Digital Libraries, East Christian Iconography, 2D and 3D Computers Graphics and Animation, Design and Modelling Cultural Heritage Objects, Preservation and Restoration of Digitalized Artifacts.