2008 Annual Conference of CIDOC
Athens, September 15 – 18, 2008
Georgios Karagiannis, Konstantinos N. Vavliakis, Stella Markantonatou, Sister Daniilia, Sophia Sotiropoulou, Maria Alexopoulou, Olga Yanoutsou, Klimis Dalianis, Thodoros Kavalieros

«EIKONOGNOSIA» AN INTEGRATED SYSTEM FOR ADVANCED RETRIEVAL OF SCIENTIFIC DATA AND METADATA OF BYZANTINE ARTWORKS USING SEMANTIC WEB TECHNOLOGIES

Georgios Karagiannis, Konstantinos N. Vavliakis, Stella Markantonatou, Sister Daniilia, Sophia Sotiropoulou, Maria Alexopoulou, Olga Yanoutsou, Klimis Dalianis, Thodoros Kavalieros.
"ORMYLIA" Foundation, Art Diagnosis Center, Institute of Speech and Language Processing, National Technical University of Athens, Project On Line.
Ormylia Foundation
Sacred Convent of the Annunciation, 63071, Chalkidiki
Ormylia
Greece
E-Mail: {g.karagiannis,ormylia}@artdiagnosis.gr
URL: www.ormyliafoundation.gr

Abstract

The documentation and analysis of Byzantine Art is an important component of the overall effort to maintain cultural heritage and contributes to learning and comprehending ones history traversal path. Efficient publishing of the multi-dimensional and multifaceted information that is necessary for the complete documentation of artworks should draw on a good organization of the data. *Eikonognosia* is a research project funded by the *Greek General Secretariat of Research and Technology (GSRT)* that aims to efficiently organize and publish detailed information about icons in the *World Wide Web.* Information derived from the analysis conducted in the *Art Diagnosis Center of Ormylia Foundation* is taken as a case study. *Eikonognosia* provides the means for organising detailed and multidimensional information about Byzantine icons in a way that is compatible to international standards (*CIDOC-CRM* - ISO 21127:2006) and allows for an easy retrieval of data with advanced semantic web technologies. The ultimate goal for *Eikonognosia* is to foster the cultural heritage community by providing an integrated framework that helps to facilitate organization, retrieval and presentation of data from the cultural heritage domain.

INTRODUCTION

The documentation and analysis of Byzantine Art aims to maintain and publish ones cultural heritage, thus proving the means for learning and comprehending one's history traversal path. The information needed for the complete documentation of artworks comes from different sources and is usually multi-dimensional and multifaceted. For this reason, publishing this information without prior understudying the users' requirements and organizing it accordingly is most of the times inefficient.

In the Art Diagnosis Center of Ormylia Foundation (OF-ADC), Byzantine artworks are analyzed with non-destructive and micro sampling analytical methods such as multispectral imaging, optical and acoustic microscopy, spectroscopic and chromatographic techniques. The information derived from this kind of analysis is enriched with descriptive, interpretative, aesthetic and technical analysis conducted at OF-ADC.

Eikonognosia provides the means for organising detailed and multidimensional information about Byzantine icons in a way that is both human oriented and compatible to international standards (*CIDOC-CRM* - ISO 21127:2006 [1]) and, eventually, allows for an easy retrieval of data with advanced semantic web technologies [2]. In this paper the overall system architecture and the design concept of *Eikonognosia* are presented.

In the next section the general architecture of the system is presented, consisting of five main modules. Each of the modules of the system, namely the conceptual module, the database layer, the ontology schema, the (semantic) web services and the presentation modules are discussed in separate sections. Furthermore a tool for instantiating the ontology schema with data from the database is presented. The paper closes with the authors' conclusions and their plans for future work.

GENERAL ARCHITECTURE

The developed system consists of five main modules. Its general architecture is depicted in Figure 1. On the first layer, data sources are organized in thematic areas in a human-oriented way. With the help of domain experts in Byzantine Iconography we categorized the data in 8 main thematic unities: Identification, Description, Interpretation, Technical, Aesthetic, and Condition data along with Multispectral Images and Spectra.

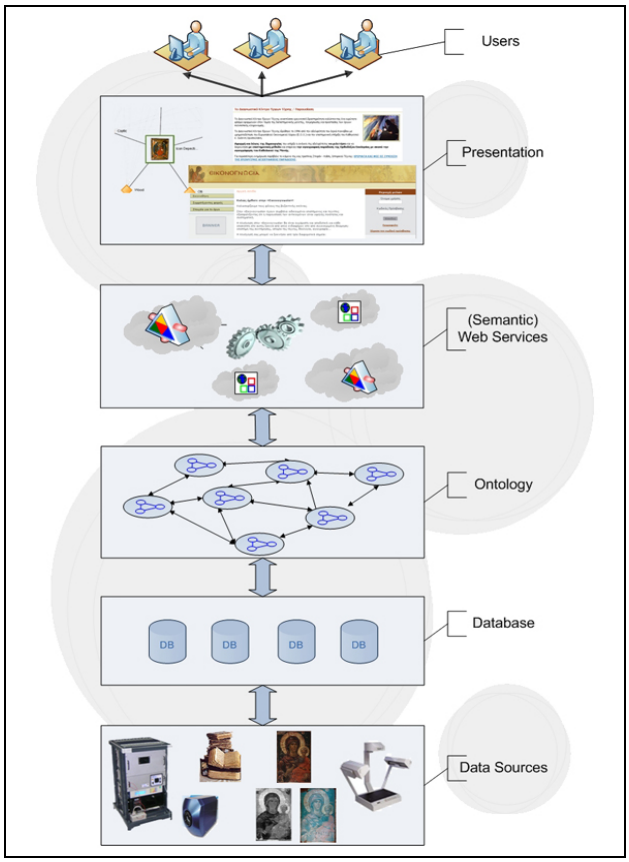

Figure 1: General Architecture of Eikonognosia

For the third module of the system an ontology schema was developed, based on *CIDOC-CRM* where the necessary classes, properties and controlled vocabularies were defined and a conceptual mapping between the database and the ontology schema was set up. Once the conceptual mapping was verified, we were able to instantiate the ontology using mapping tools that transformed the relational data to semantic data and allowing for the semantic querying of the relational database.

The next layer consists of all the (semantic) web services that provide the methods to efficiently search, access and retrieve all the documentation of an art object using the *WWW*. Some of the developed web services may be customized to specific users' profiles, while other may be generic. These services provide the ground for the next module.

The final layer of the system is the presentation module were all the information is presented to the final users via a web interface. In this way *Eikonognosia's* web site provides a wealth of information, optimized according to the experts' and/or simple users' requirements.

ORGANIZATION OF DATA

As already mentioned, the information needed for the complete documentation of artworks is multi-dimensional, multifaceted and extremely difficult to be presented in a way that is both simple and revelatory of the nature of the object. In our effort to face this problem we opted for organising the data in a human-oriented way, without neglecting any necessary information. In this effort the help of domain experts on Byzantine Art Images was crucial. We categorized out data into 8 main thematic unities, which are shortly presented in Table 1.

The first thematic unity consists of the Identification Data of the artwork. In this unity data about the title, the main theme and style, the creator and school, the provenance and the actual owner, the type and dimensions of the object are included. The purpose of these data is to provide general information about the artwork and to uniquely identify it.

Data		Description of Data
Level 1	**Level 2**	-
Identification		Data that define the artwork

Description Data	**Title**	The title of the object
	Theme	The main theme of the object
	Dimensions	Height, width and depth of the object
		Descriptive information about the image's characteristics
	Theme Description	Description of the theme (general and specific)
	Prototype Objects	Objects that are regarded as prototypes of the image
	Inscriptions	Description of the inscriptions (title and signatures)
Interpretation Data		Interpretative meaning of everything aspect of the image that can contain religious information
	Theme Interpretation	Interpretation of the main theme
	Inscriptions	Interpretation of the inscriptions of the artwork
Technical Data		Description of the technique, materials and tools
	Support Material	Analysis of the support material
	Stratigraphy	Analysis of the stratigraphy layers and their pigments
Aesthetic Data		Description of the aesthetic of the image
	Composition/Perspective	Description of the composition and perspective of the artwork
	Drawing Style	Description of the drawing style
	Color	Description of the palette used by the painter
Condition Data		Alterations and damages that the image has suffered, along with data about the restoration process
	Paint Layer Damage	Description of the damage in the paint layers
	Overpaintings	Description of the overpaintings
Images	-	Multispectral images justifying the descriptions
Spectra	-	Spectra justifying the descriptions and analysis of condition

Table 1: The main thematic unities of the data

The next unity comprises the Descriptive Information about the artwork. For example the description of the theme and its visual representation on the icon is provided, along with possible inscriptions or entities/objects appearing on the painting. Furthermore other related artworks that may constitute the prototype or variations of the object are presented. This unity aims at both describing in detail the icon and presenting its interconnections with other icons according to theme/creator/style/period.

In the third thematic area Interpretative Data are provided for almost all the descriptive information of the second unity. In Byzantine Iconography almost every element of the art object has either a specific theological meaning or a liturgical tradition or an historical background. This thematic unity aims at interpreting, where possible, the elements and the attributes of the art object.

The fourth thematic unity contains technical data. They analyze the material structure of the artwork (the support and its preparation for painting, the stratified paint and varnish layers, the pigments, binding materials and the general technique of the artwork.

In the fifth unity aesthetic data are given. The aesthetic data include useful information about the composition, drawing and the colors the artist used. Moreover the influences of the artist are documented. Most of these data derive from color measures as well as from close observation of the object itself and its technical images such as infrared or X-Ray images by domain experts.

The thematic area of condition contains information about the current condition of the artwork and its materials in width and in depth. The over-paintings and depositions are analyzed along with the induced damage in the paint layers, the varnishes and the support material.

The last two thematic unities of images and spectra provide information that complements and documents with visual or analytical evidence all other thematic unities as it is the case that important conclusions can be derived through close observation on multispectral images and the pattern of spectra. For this reason, together with the description of an element of the artwork the appropriate image is always presented. Likewise the conclusions derived from spectral analysis are given together with the appropriate spectra.

DATABASE SCHEMA AND ITS GRAPHICAL INTERFACE

The second module of the system stores all the information derived with (non)-destructive methods, along with the analysis of the domain experts of the first layer, in a relational database, developed at *Oracle*^{TM} RDBMS. A complex relational schema was developed based on the analysis conducted for the organization of the data. The database schema is fully compatible with the organized data coming from the first module. The digital images are stored in the database using the *Oracle*^{TM} *InterMedia* [3] technology. The database consists of more than 250 tables, allowing the organization and storage of all types of data necessary for the complete documentation of Byzantine Art Objects.

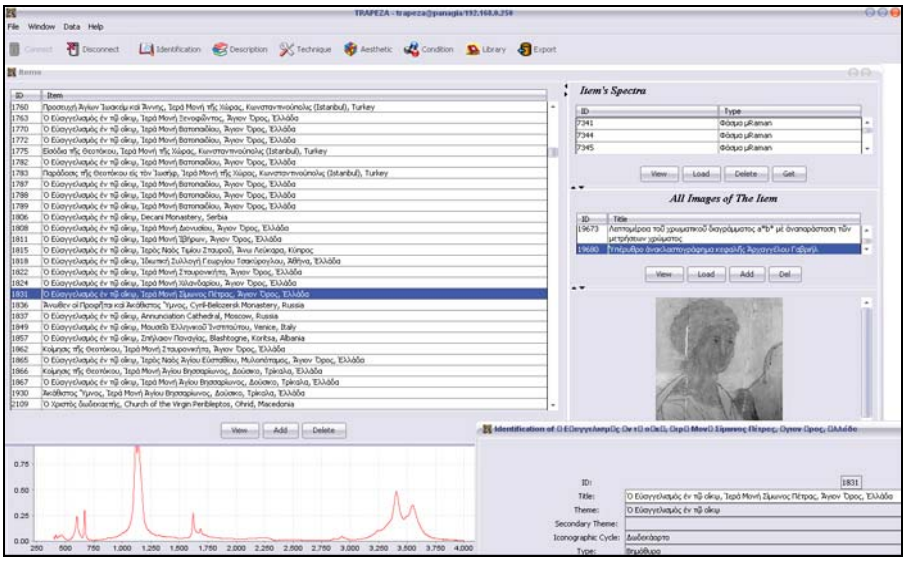

Figure 2: Interface of the database

The complicated schema of the database is due to the wealth of information that is recorded in *OF-ADC* as well as to the fact that almost all data are accompanied with digital images and spectral information supporting the text. Moreover significant effort was paid to modeling the data and create as many controlled vocabularies as possible in order to organize better the information and create dictionaries of concepts.

Along with the database schema a graphical interface was developed providing navigation and insert/update capabilities throughout the database (Figure 2). The

graphical interface was built in Java language. Besides viewing or adding text, the developed graphical interface allows the manipulation of image and spectra. The user can add/delete/update image and spectra files and the appropriate interfaces with zooming capabilities are provided. The user can also save any type of document or file in the database.

AN ONTOLOGY FOR BYZANTINE ART OBJECTS

The next module of the system consists of ontology compatible with *CIDOC-CRM*. The ontology was enriched, as a first approach, with the necessary classes, properties and controlled vocabularies, and a mapping between the database and the ontology schema was set up that enables semantic querying of the relational database.

The adequate documentation of all aspects of Byzantine icons, which are both liturgical objects of the Orthodox religion and works of art, crucially also, relies on the detailed documentation of their material parts. As a first, under discussion approach, an extension to *CIDOC-CRM* in order to capture a large set of multimodal data (textual, visual and, to a lesser extent, audio) organized into a dense network of relations was proposed. As already explained, these data provide detailed information about the identity, technique, condition, conservation, aesthetics, thematic content, as well as the signifier function of icons within the Christian Orthodox community. In particular, the signifier function of icons formed the basic organisation principle for the ontology.

Icons are portable paintings with a major liturgical role in the Orthodox religion. These properties set icons apart from other objects. We define the class E22.x Liturgical Object for the set of objects that play a liturgical role in some religion (a rigid property [4]) and E22.xy Icon (a subclass of E22.x Liturgical Object).

In order to document the fact that a rather piece-wise correspondence between the meaning of icons and their material and visual features has been established in the literature on Byzantine icon interpretation [6] we propose a set of subclasses of E22.

Man_Made_Object for the material parts of the icon: carrier, preparation layer, paint layer, metal parts and varnishes. Each material part can be decomposed to the necessary sub-parts indexed by an appropriate name (subtypes of E55). For instance, the paint layer can be decomposed to sections indexed by names such as 'μαφόριον' and 'κάμπος'.

The icon and its parts can be related with instances of E36 Visual Item indexed by their name, e.g. 'μαφόριον'. We distinguish between the actual appearance of the icon (instance of E36) and "Icon Composition", that is, the visual impression created by the paint layer when the icon was brand new. E38.x Icon Composition (subclass of E38 Visual Image) is related to the icon and the paint layer and is indexed by iconographic type, which is, in turn, indexed by iconographic cycle – both established icon classification systems. Descriptions of the entities of these classes are provided with the ***P3 has note*** property.

P129 is about E28 Conceptual object relates the visual objects with their meaning (instances of E28). We define the property E22.xy Icon $\boldsymbol{P_N}$ *carries symbolic (rather semantic) content* E28 Conceptual Object as a shortcut between the material (parts of the) icon and their meaning.

The icon can be related with its meaning directly if the meaning is shared by all the Orthodox community, or via an E13.x Interpretation Assignment event (subclass of E13 Attribute Assignment) if the meaning is the intellectual product (rather interpretation) of an identified individual.

The proposed documentation of the material (parts of the) icon and the corresponding visual objects allows for a documentation of interpretations and interpretation events that closely follows the way icon interpreters work.

INSTANTIATING THE ONTOLOGY SCHEMA

The connection of the database with the ontology can be done in a number of ways. In *OF-ADC* various approaches have been applied, such as using the *D2RQ* engine [6] to transform semantic queries to *SQL* queries, or using an evolution of the *Eikonomia* [7] engine to copy the data of the database to the ontology. Each method has its own pros and cons, for this reason a tool was developed that allows the user to select the method best suiting her needs.

The general architecture of the transforming mechanism is depicted in Figure 3. The transformation algorithm uses the Jena [8] or the Sesame [10] semantic frameworks to instantiate an ontology schema with data from a relational database. The outcome of this operation is an ontology containing the data from both the original ontology and the database.

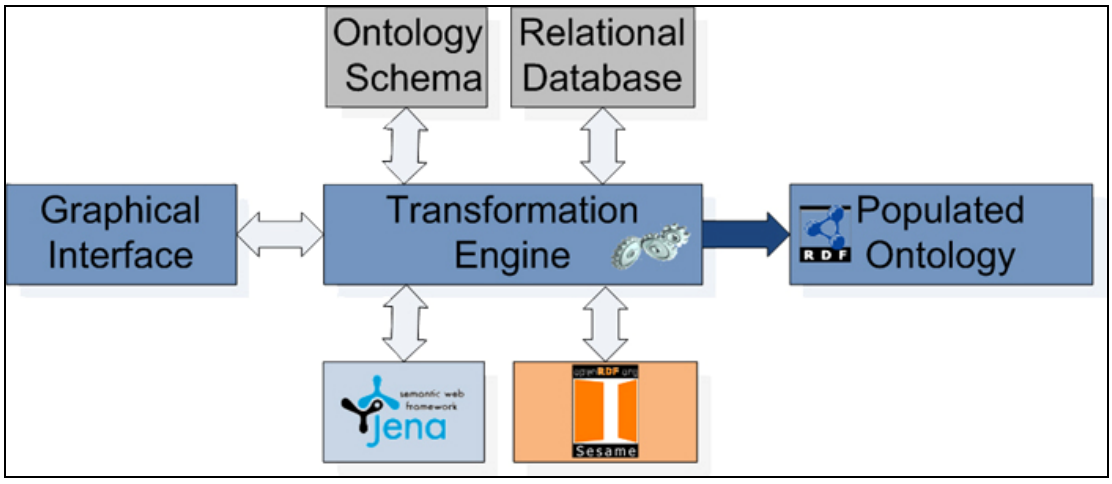

Figure 3: Architecture of the Transformation Algorithm

The *Eikonomia* algorithm is built around two main steps. In the first one the user has to set up the mappings for the instantiation of the ontology schema and in the second she has to define the object or datatype properties applying to the newly created instances. Various helping mechanisms have been developed in order to reduce typing errors and improve productivity.

In the case of the *D2RQ* algorithm, users have to define ClassMaps and PropertyBridges in order to map the database to the ontology schema. A suitable

interface was developed that allows the user to easily create the mappings with the *D2RQ* vocabulary. The interface developed supports drag-n-drop, auto complete and other checking mechanisms in order to reduce typing errors and provide a user-friendly environment.

(SEMANTIC) WEB SERVICES & PRESENTATION

In the last sections we have described the logical organization of the data, their storage in the respective database schema with the help of a graphical user interface and the development of a *CIDOC-CRM* compatible ontology that allows semantic access of the data. So far, we have created a rich semantic repository, which allows us to process all types of data in a number of ways. Next, we discuss the developing of appropriate methods for efficiently searching, viewing and interchanging data by employing Semantic Web technologies.

The final two architectural levels were based on the *e.Centric* platform [10] developed by *Exodus*. A mock-up of the web site is depicted in Figure 4. The *Exodus e.Centric* platform provides the framework for the development of specific modules, which materialise the desired functionality within the *Eikonognosia* project. The platform has been designed to be easily expandable and uses latest generation standards such as *XML*, *XSLT*, *XML* Schema, *WEBDAV*, *ICE*, *RSS* and Web Services, while offering full functionality via the use of the *HTTP* protocol. *ASP* system expansions allow an *e.Centric* Administration site to host multiple projects. The platform's architecture ensures that it can be used with a wide variety of devices (web farms, automatic content replication, scalable session management), which allows work load to be allocated and instances of single point of failure avoided. The platform has also been designed to function with the majority of database systems.

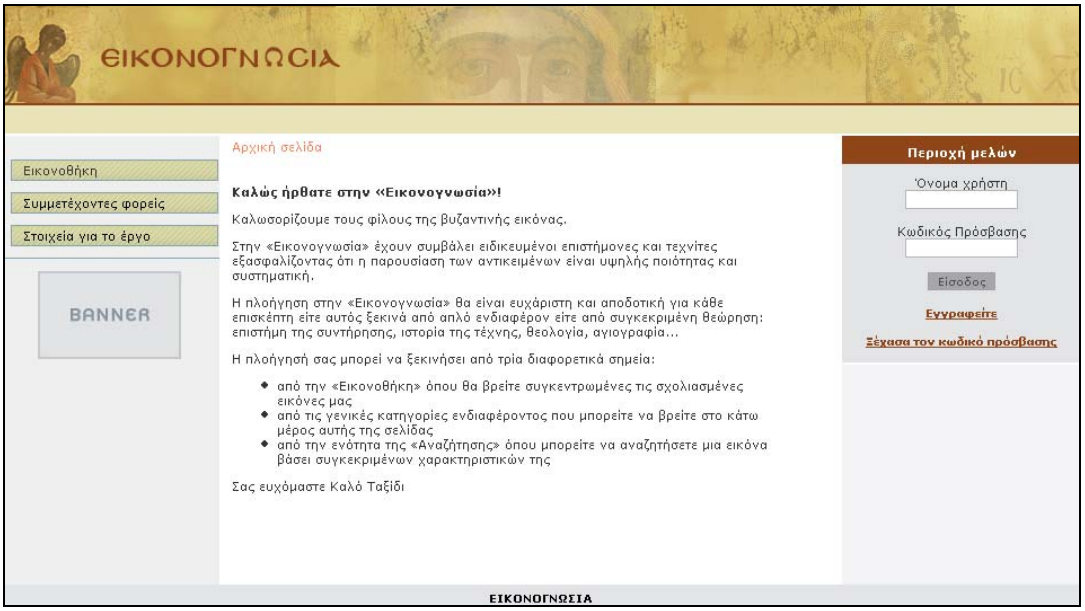

Figure 4: Mock-up of the Final Site

Generally speaking, the Exodus *e.Centric* platform adopts an object-based approach and allows for the definition of smart modules by and for users. It can also share an interface with existing operating systems. *e.Centric* has been used as the basis for the implementation of the site of *Eikonognosia*. Based on the users' requirements analysis and the defined specifications, all the pages have been produced and the user interface has been put in place. The user/visitor of the site may easily navigate through the different sections of the site and browse the available information. All the information is stored in external sources (*Oracle*^{TM} database) and the portal communicates with the source in order to retrieve and display the requested information.

On top of that, the so-called "*Taxonomy Management Module*" was developed from scratch, in order to support the need for multiple categories per object. The new module supports, with a very friendly user interface, the creation of a taxonomy in the form of a "*forest*" of trees, which allows the visual representation of the relations between data units. Within this forest new nodes (i.e. categories) may be added, modified or deleted. If a taxonomy already exists in the form of *RDF* standard, the module allows both its automatic import and the proper generation of all the corresponding nodes. This taxonomy may be used for "*tagging*" objects with additional information.

CONCLUSIONS & FUTURE WORK

Eikonognosia aspires to be the first multimedia, multilingual and user-friendly portal using semantic web technologies that will provide users a wide range of information about the Byzantine Art domain and will constitute a place of reference to everyone interested in Byzantine Hagiography. Through *Eikonognosia*, the domain expert is expected to easily access and manage information on artworks and their interdependencies from a semantic perspective. We hope that the *Eikonognosia* will provide to the layman a generic description of the artwork with the option to look for more detailed and technical information whenever he wishes.

The design of *Eikonognosia* is such that the underlying knowledge model of the system may be expanded to comprise additional forms of cultural heritage data. Moreover the use of semantic web technologies, along with the use of the *ISO* standard CIDOC-CRM, provides *OF-ADC* the potentiality for data and service interoperability.

ACKNOWLEDGEMENTS

This work has been supported by the *Eikonognosia* project, entitled "Development of a knowledge management tool for accessing scientific data and metadata of Byzantine Art Objects using advanced semantic web tools for sorting, indexing, retrieval and presentation of information (*Eikonognosia*)". The project was funded by the Greek General Secretariat of Research and Technology (GSRT).

REFERENCES

[1] N. Crofts, M. Doerr, Tony Gill, Stephen Stead, Matthew Stiff (editors), Definition of the CIDOC Conceptual Reference Model, October 2006. URL: http://cidoc.ics.forth.gr/definition_cidoc.html, accessed at 29 May 2008

[2] T. Berners-Lee, J. Hendler, and O. Lassila. The SemanticWeb. Scientific American, May 2001.

[3] Oracle Inc., Oracle InterMedia User's Guide, 10g Release 2, June 2005

[4] N. Guarino and Christopher Welty. 2002. "Evaluating Ontological Decisions with ONTOCLEAN". Communications of the ACM, February 2002, Vol. 45 No.2, pp. 61-65

[5] L. Ouspensky and V. Lossky. 1999. The meaning of Icons. St Vladimir's Seminary Press, Crestwood, New York 10707

[6] R. Temple. 1990. Icons and the mystical origins of Christianity. Element Books, Longmead, Shaftesbury, Dorset

[7] C. Bizer, A.Seaborne, D2rq - treating non-rdf databases as virtual rdf graphs,in: ISWC2004, 2004, URL: http://www4.wiwiss.fu-berlin.de/bizer/d2rq/; accessed at 29 May 2008

[8] K. Vavliakis, A. Symeonidis, G.Karagiannis, P. Mitkas, Eikonomia – An Integrated Semantically Aware Tool for Description and Retrieval of Byzantine Art Information, ICTAI 2007, October 29-31, Patras

[9] J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, K. Wilkinson, Jena: Implementing the semantic web recommendations, Tech. rep., HP (2003). URL http://jena.sourceforge.net/ ; accessed at 29 May 2008

[10] J. Broekstra, A. Kampman, F. van Harmelen, Sesame: A generic architecture for storing and querying rdf and rdf schema, in: ISWC2002, Lecture Notes in Computer Science, Springer, 2002. URL http://www.openrdf.org/ ; accessed at 29 May 2008

[11] e.Centric platform, Exodus, URL: http://www.exodus.gr/%287005011066027299%29/ecPortal.asp?id=181&nt=55&lang=2, accessed at 29 May 2008