BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **9**, No 1

Sofia • 2009

Ontology "Bulgarian Iconographical Objects" – Creation and Experimental Use*

Kamenka Staykova, Danail Dochev

Institute of Information Technologies, 1113 Sofia
E-mails: {staykova, dochev}@iinf.bas.bg

Abstract: *The paper presents the phases of designing and implementing the subject domain ontology "Bulgarian Iconographical Objects" for the relevant tasks of the project LOGOS* ***"****Knowledge-on-Demand for Ubiquitous Learning". The results of ontology involvement in annotation and search processes of digital learning materials are commented. Further improvements of these processes are discussed together with the relevant conceptual extensions of the ontology.*

Keywords: *Ontologies, knowledge representation, conceptual graphs.*

1. Introduction

The ontology "Bulgarian Iconographical Objects" was created under the FP6 project LOGOS **"**Knowledge-on-Demand for Ubiquitous Learning" **http://logosproject.com**. The project deals with accessing existing large-scale repositories of digitized information and facilitating their transformation into learning content, considering also possibilities for its cross-media delivery. The overall aim of the project is to contribute to the adequate enhancing and facilitating the knowledge building during eLearning processes. Ontologies and their processing were included in this project in order to realize content- and context-sensitive search in large digital repositories with considerably less explicit input information from the users side. Ontology "Bulgarian Iconographical Objects", OBIO for short, is one of the LOGOS learning domains' ontologies and contains the main concepts in the subject field "Bulgarian Iconography". OBIO targets semantic annotation and search of digital objects, which set up the creation of

*This work was partially supported by IIT-BAS Project № 010079 "Methods and Tools for Processing Semantic Information".

courseware materials in the learning domain of Bulgarian Iconography, to be used in several learning areas (iconography, arts, history, culturology, theology, etc.).

The design and development of OBIO followed the classical phases of constructing knowledge applications as ontologies. According to [4, 6] the construction of an ontology consists of work on three main levels: knowledge level, symbol level and implementation level. Specification of ontology purpose, usage, scope, degree of formality and data collection through various elicitation methods are needed steps in the *knowledge level.* Conceptualization of domain terms, resulting in a preliminary ontology followed by possible integration with other ontologies covers the *symbol level.* The *implementation level* of knowledge applications development consists of formalization in an ontology language and evaluation of completeness, consistency and redundancy of created ontological structure.

The main part of OBIO domain vocabulary was determined at the knowledge level. Main ontology concepts, the relations among them and the taxonomies of concepts and relations were specified. The knowledge elicitation was an iterative process of interviews and brainstorming sessions with domain experts and developers of the multimedia library "Virtual Encyclopaedia of the Bulgarian Iconography" [5] from the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences. The Conceptual reference model of the International committee for DOCumentation (CIDOC) was chosen as a top level ontology. It is developed by the **CIDOC Documentation Standards Working Group** of International Council of Museums [3] and is a standard for description of art collections and cultural heritage artifacts.

The work on symbol level concerns formalization of the knowledge by using Conceptual Graphs formalism. According to the Conceptual Graphs terminology [2], each ontology is presented by a construction called *support* and contains taxonomy over a vocabulary of *Concept Types* and taxonomy over a vocabulary of *Relation Types*. *Individuals* could also be provided. Each statement of knowledge is given by a *Conceptual Graph.* A simple *Conceptual Graph* defined over a *support* is a graph containing two kinds of nodes: concept nodes, which represent entities and relation nodes, which connect these entities. Homomorphism (or projection) is the main notion of *Conceptual Graphs* processing. While *Conceptual Graphs* represent formalized knowledge, projection is the core operation of a knowledge management system, which enables the enforcement of explicit knowledge in form of *Rules*, *Constraints*, etc.

The Ontology Management Tool of LOGOS Authoring Studio was used at the implementation level. The Ontology Management Tool is based on the instrumental system CoGui **http://www.lirmm.fr/cogui/**, developed at the Montpellier Laboratory of Computer Science, Robotics, and Microelectronics (LIRMM). It facilitates the work of ontology engineers, who construct ontologies in conceptual graphs by all standard functions: visualization, easy knowledge insertion and deletion, ontology verification, etc.

In the following Section 2 the OBIO ontology is presented by itself after its creation and first employment within the LOGOS project. Section 3 gives more

details on the very process of OBIO usage, namely the annotation process launched into the area of digital iconographical images of Orthodox icons. Concerning the ontologies' life-cycles, Section 4 speaks about future enlargement of OBIO and possible scenarios of its usage for different research experiments. Our conclusions obtained from the work done for now are presented in Section 5.

2. Presentation of OBIO ontology

2.1. OBIO Life-Cycle

Currently the society of knowledge engineers is seriously concerned by using ontologies for large-scale semantic applications in distributed organization, improving the capability to handle multiple networked ontologies that are created collaboratively within a particular context, and might be highly dynamic and constantly evolving[2]. Ontologies (or ontology networks) life-cycle schemes are in the centre of attention. Five different models of ontologies' life-cycles are outlined: *Waterfall*, *Iterative Model*, *Incremental Model*, *Evolving Prototyping* and *Spiral model*. The choice of the particular model to follow depends on the dynamics of requirements to the future ontology, on the need to be produced and implemented all intermediate results or not, on the need to include risk-control activities to the development process. 55 different activities of creating ontologies (networks) are identified. Some of the activities are mandatory like *Ontology Specification, Ontology Conceptualization, Ontology Formalization;* others are required only in case they are applicable in the particular case of ontology building (*Ontology Translation, Ontology Localization, Ontology Alignment*, etc.).[3]

The particular ontology life-cycle is a result of selecting the proper ontology life-cycle model first, and then – selecting the activities, which are recommended and applicable. In our work to create OBIO ontology the *Incremental Ontology Life-Cycle Model* was chosen as consequence of *Ontology Feasibility Study*. After *Ontology Search* and *Ontology Selection* activities, the Conceptual Reference Model of CIDOC was chosen as upper level conceptual guide. *Ontology Specialization,* or in particular specializing CIDOC CRM concepts to get the lower conceptual level of OBIO was recognized as important activity of the new ontology development. Two important decisions were taken at the *Scheduling* process. First, the *Translation* of CIDOC CRM to the formalism of Conceptual Graphs was planned to be combined with *Ontology Alignment* at later phase of the development process; Second, *Ontology Forward Engineering* was performed with activities of *Knowledge Acquisition*, *Ontology Specification*, *Ontology Conceptualization*, *Ontology Formalization*, *Ontology Implementation* in conceptual graphs, see Fig. 1 below. OBIO Life-cycle contains also activities of *Ontology Localization*, concerning the Bulgarian and English contexts of OBIO, *Ontology Update* and *Ontology Upgrade*, *Ontology Validation*, *Ontology Verification*, *Ontology Assessment* and *Ontology Extention* in knowledge constructions of time and space.

[2] For more details see NEON projects' site **http://www.neon-project.org**.
[3] All the descriptions of the 55 activities concerning ontologies life-cycle could be seen at **http://www.neon-project.org/web-content/images/Publications/neonglossaryofactivities.pdf**.

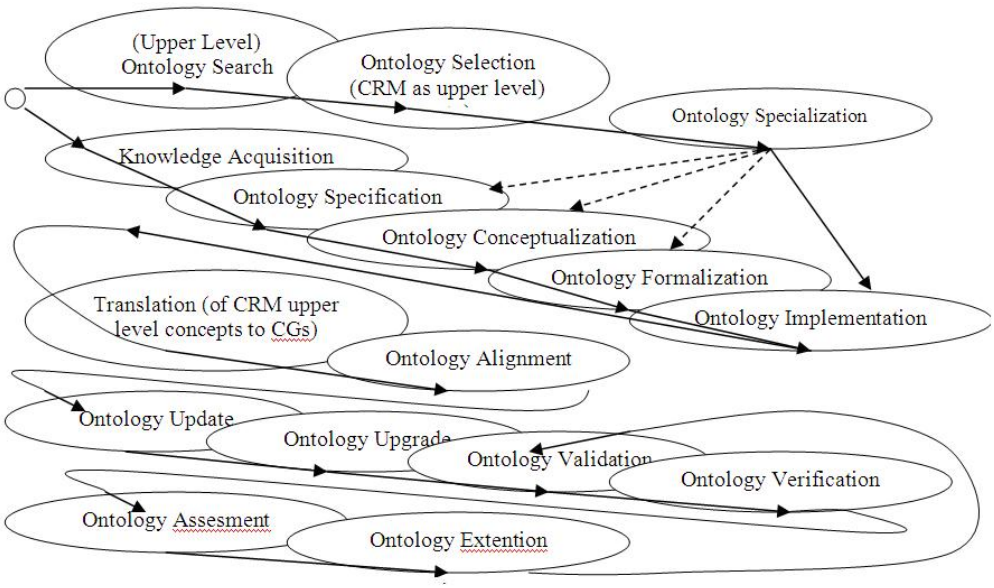

Fig. 1. OBIO Life-Cycle

The ontology OBIO as a result of the *Implementation activity* is presented here in the style and outlook of conceptual graphs, with its vocabulary (2.2), and its formalized knowledge in graphs (2.3).

2.2. OBIO vocabulary

Concept Types

The OBIO hierarchy of *Concept Types* includes 62 notions. Some of them show the specificity of "The Domain" like *Icon*, *Manuscript Book*, *Iconographical Image*, *Iconographical School*, *Iconographer*, etc. and others are more general notions about time and place, which are also needed for targeted knowledge representation.

Relation Types

The hierarchy of *Relation Types* consists of 32 binary relations like, for example:

Iconographical object	***has current location***	*Place*
Iconographer	***had worked during***	*Time-Span*
Iconographical object	***shows image***	*Iconographical Image*

The organization of both *Concept Types'* and *Relation Types'* hierarchies follows the logic of notions of upper levels defined by CIDOC Conceptual Reference Model.

Individuals

In general, each *Individual* presented in the ontology is an instance of a given *Concept Type*. During the first round of ontology employment, 212 *Individuals* were introduced in OBIO. They are instances of different *Concept Types* like *Iconographer*, *Iconographical School*, *Place*, etc.

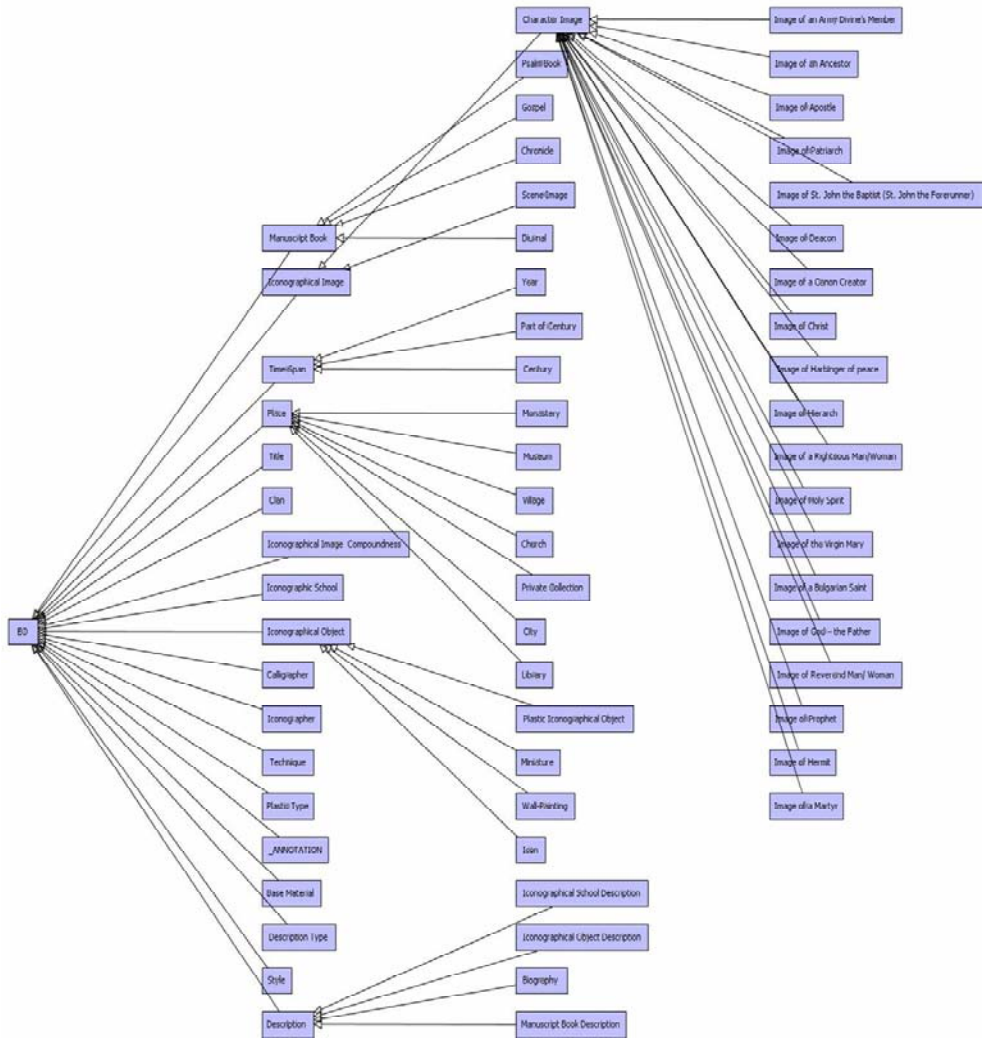

Fig. 2. Graphical presentation of OBIO Concepts Types hierarchy

Nesting Types

Nesting types are used in knowledge representation with conceptual graphs to express the complexity of some entities. The concepts which comprise (or could belong to) several already defined *Concept Types* are defined explicitly as *Nesting Types*. *Nesting Types* present one of the main differences between Conceptual Graphs logical formalism and the most popular knowledge representation languages now like OWL, RDF, FLogic, OCML, etc.

Within OBIO ontology *Nesting Types* are used to facilitate the ontology involvement in the process of digital images annotation. The root of OBIO nesting hierarchy is called *Annotation View* and each nesting type defined there is meant to connect a real annotated object (text file, HTML file, digital picture, etc.) to a corresponding OBIO knowledge construction, or pattern of knowledge, presented by a particular *Pattern Graph*.

2.3. OBIO Domain Knowledge

Facts

Facts are simple graphs, which contain *Individuals* and express factual statements about them. The OBIO first version contains 27 Facts.

Fig. 3. A fact from OBIO

Rules

Rules are posed in the ontology to spare the insertion of factual statements and when the needed conclusions could be derived from some already existing knowledge constructions. OBIO contains five Rules. For example,

when an Iconographer *is a member of* a Clan
and a Clan *belongs to* a given Iconographical School,
then it could be automatically derived that
the Iconographer *is representative* of that given Iconographical School.

Prototypic Graphs

Prototypic Graphs present knowledge focused on a particular *Concept Type*, which integrates some of the connected to that concept other *Concept Types* through the relations of the ontology. There are defined within the OBIO *Prototypic Graphs* for *Iconographer*, *Iconographic School*, *Iconographical Object* and *Manuscript Book*.

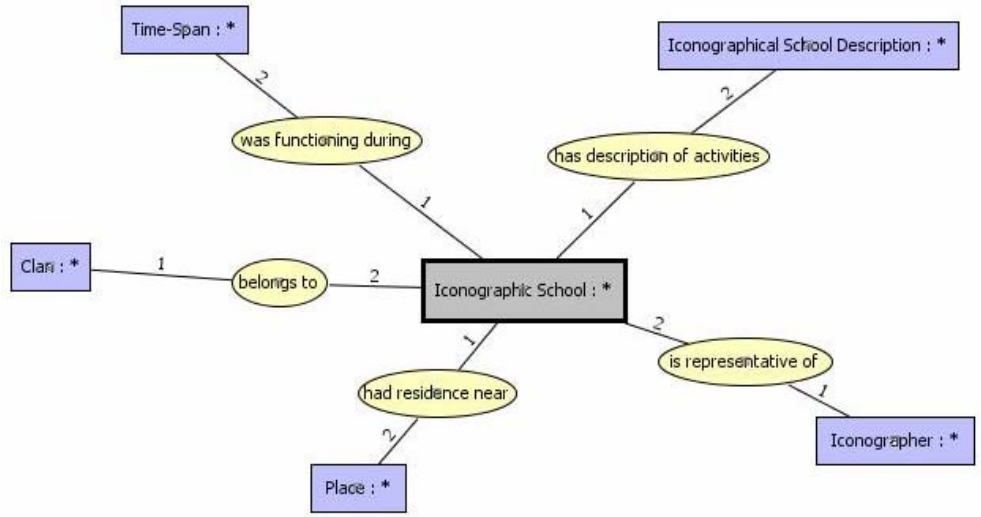

Fig. 4. The Prototypic Graph of the concept “Iconographic School”

Individual Graphs

Through their type the newly introduced *Individuals* in OBIO could be connected to the already defined specific knowledge prototypes called *Individual Graphs*, and this connection makes easier the process of data insertion.

Pattern Graphs Currently OBIO contains 8 *Pattern Graphs*, which are connected to the defined *Annotation Views* and ensure the practical usage of ontological knowledge during the process of annotation.

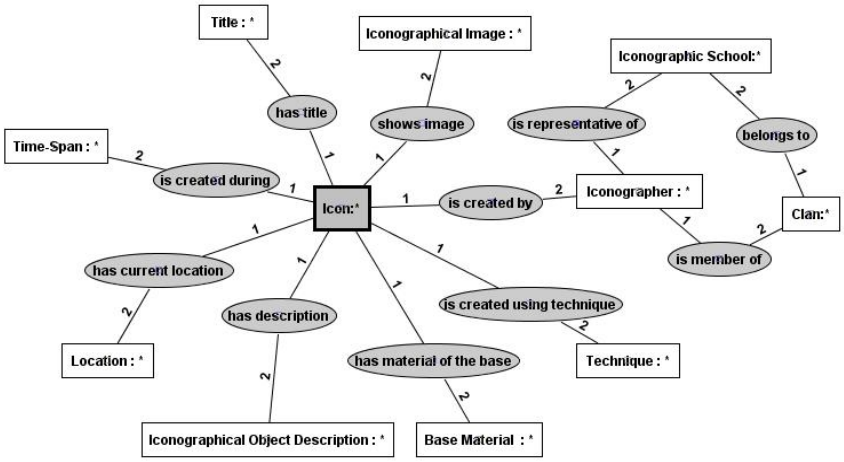

Fig. 5. Pattern Graph of the concept “Icon”

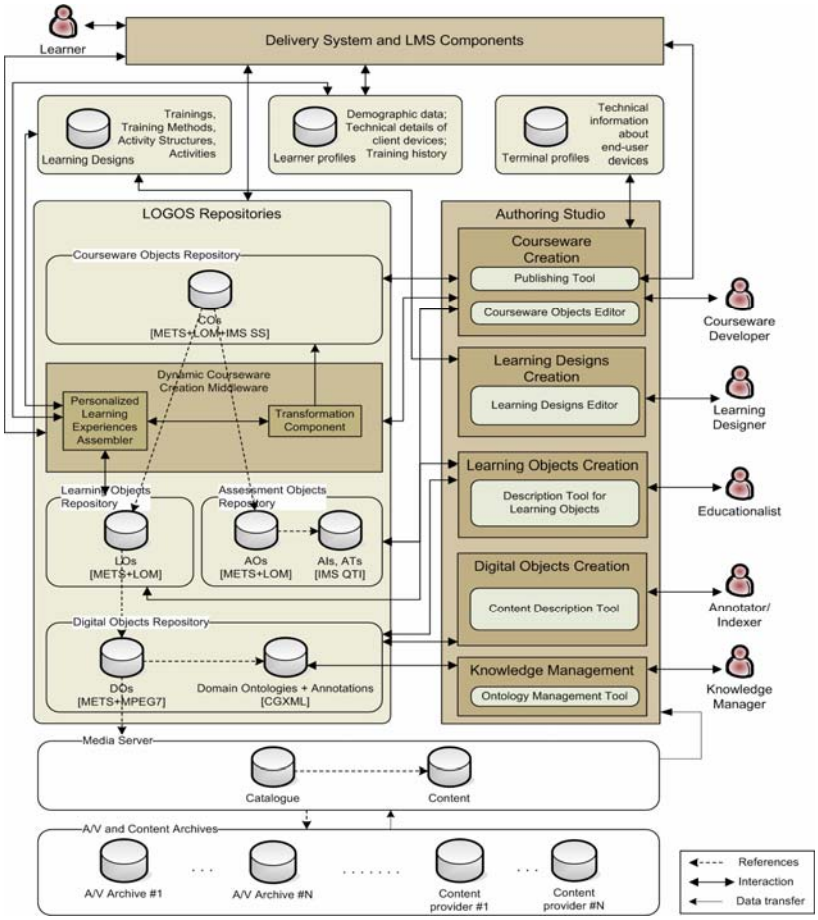

Fig. 6. LOGOS Authoring Studio: Overall Architecture

3. Annotating digital images of orthodox icons using OBIO within the LOGOS studio

Once implemented the OBIO ontology was involved in the annotation process within the experimental subject domain of LOGOS Authoring Studio 'Bulgarian Iconography'. The Fig. 6 presents the architecture of the LOGOS Authoring Studio [1]. Below is considered the first phase of the preparation of learning materials when bottom-up direction is followed and learning courses are created from raw digital materials. This process begins with annotation of *Media Objects* and creation of *Digital Objects*, which are the core bricks to build learning courses.

When the subject domain is chosen, the real process of semantic annotation supported by the LOGOS Authoring Studio aims to connect a Media Object (picture, document, etc.) from the Media Objects Repository with a knowledge construction of particular domain ontology and to store the record of this connection, or annotation, to the Digital Objects Repository. In case with OBIO in use, the chosen domain is Bulgarian Iconography, the Media Objects Repository contains digital pictures of icons and other documents of relevant information, and these files have to be annotated against OBIO ontology. In fact, each Media Object is connected to a knowledge construction of the OBIO ontology, which is accessed through a particular *Annotation View*. For example, if we have a digital picture of an icon in the Media Objects Repository, we are expected to annotate this media object against the OBIO ontology using *Annotation View to Icon.*

The annotation is performed with a tool of the LOGOS Authoring Studio called **Content Description Tool**, which is designed and implemented by a team of researchers from Institut National de l'Audiovisuel (INA), France. The terminology of annotation process contains the term *Global Annotation*, which meaning is "description of the chosen media object as a whole": one media object receives an annotation and corresponds to one digital object in the Digital Objects Repository. When the annotated media object is an image (digital picture of the public formats .jpeg, .bmp, etc.), it is possible some rectangular spatial sections of the image to be also annotated by their own besides the *Global Annotation* of the whole picture. Each one of the section-annotations creates a new Digital Object, because it provides a connection between an image and a corresponding knowledge construction; this way the description of the whole image contains several Digital Objects. It is of great importance that the ontological knowledge constructions used during the annotation process (i.e. the available *Annotation Views*) have to be well organized and carefully defined. OBIO ontology contains formalized knowledge to be annotated digital pictures of icons like this shown on Fig. 7 bellow. The *Iconographical Image* "Deisis with the Apostles"[4] is a *Compound Image* and besides the central *Image of Christ* it presents portraits or *Images of Apostles*, ordered according to particular canonical rules.

[4] The word "Deisis" comes from a Greek word meaning "prayer" or "intercession". In iconographic language it represents a group of three persons: Christ, seated in majesty in the center, with his Mother to his right and John the Forerunner to his left.

The working loop of an annotating session of the LOGOS Content Description Tool begins with loading and visualizing the chosen media object, in order to have visual control on what the annotated media object is by itself. Then a form-sheet of this media object has to be created (or loaded if the form-sheet already exists). Then the *Global Annotation* for the media object should be fulfilled, which involves the semantic model, namely the particular ontology with its knowledge patterns. At the next step, if partial annotations have to be introduced, they are created by selecting a spatial box from the uploaded image (section), and associating this section with an appropriate knowledge construction of the ontology. Again, the knowledge pattern is filled with the specific information about the content of the selected section. The process of annotating all needed sections is an iterative process which ends with saving the whole description (the *Global Annotation* and all section annotations) within Digital Objects Repository.

Fig. 7. Picture of an icon “Deisis with the Apostles” (unknown painter from Etropole, beginning of 19th century)

The annotator has some freedom while fulfilling and completing the global or partial knowledge patterns. During the annotation process the items which are not relevant could be suppressed, or some knowledge constructions could be duplicated, which alters the basic knowledge pattern (*Annotation View*). For each concept of the resulting conceptual graph the annotator could specialize the involved concepts using *Concept Types* hierarchy of the ontology, or the concepts could be “individualized” with existing *Individuals.* New *Individuals* could also be

inserted using the *Individuals' Server*, see Fig. 8 bellow. *Individuals* are tightly linked to the ontology (a particular *Individual* belongs to a particular ontological *Concept Type*), but they are stored and accessed independently by the *Individuals' Server*, because they are the only ontology terms which could be defined "on the fly" by the annotators.

Once completed with all the descriptions a Digital Object (annotation) is supplied with three links:

– a link to the corresponding Media Object; it can be the simple URL of the media in the case of a global annotation, or the URL may be completed by a spatial selection (a region coordinates). This kind of link is called *Anchor*.

– a link to a form-sheet; this form is useful to give some free textual information during the annotation, most often the title of the Media Object and a short description in a form of abstract (to be in help for free-text search, for example). This form is also used to manage the status of the annotation, and to record some implicit information tracked by the software (the current author, previous authors, technical data about the content, etc.).

– a link to the conceptual graph, which presents the knowledge construction used for the semantic annotation; this shows the connection to the pre-defined semantic model (the ontology) reflecting an editorial will or information about the Media Object from particular *semantic* point of view, in order the Media Object to be easily retrieved later.

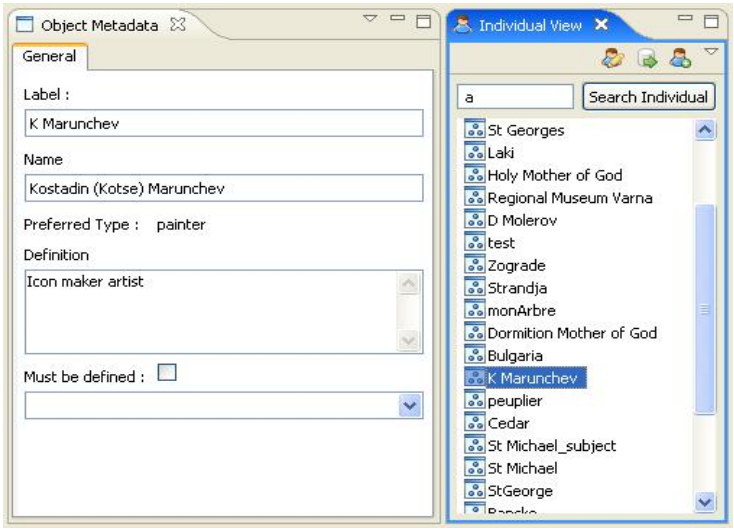

Fig. 8. Screen-shot of CDT interface: working with Individuals' Server

More than 100 Digital Objects with their corresponding descriptions were created during the internal check and functionality evaluation of the LOGOS Authoring Studio doing tests within the particular subject domain of Bulgarian Iconography and OBIO. The annotation process offered by the Content Description Tool was tried with all its facilities. Considering the experience in that particular subject domain, it could be claimed that the tool provides all the functionalities

needed for digital pictures annotation, which is the target of the particular project state. It is also clear, that being a part of a research environment, the Content Description Tool can not give very sophisticated support to massive annotation work of many annotators working in parallel; or can not give special on-the-flight support to the annotators offering, for example, the nearest knowledge patterns. The Content Description Tool of LOGOS Authoring Studio works stable and supplies the annotators with comfortable and user-friendly interface; from a research perspective it realizes the basics for further research experiments with the annotation process.

4. Future experiments

Often the annotators' work is slow and monotonous, especially when semantically similar data has to be inserted many times, and the same patterns have to be chosen and fulfilled repeatedly. The main task of the annotation systems is to give annotators some clues in these cases, and to offer "the nearest" patterns from the set of already fulfilled and stored knowledge patterns. The idea of dynamically supported process of annotation is the leading line of the further experiments with the Content Description Tool of LOGOS Authoring Studio and OBIO Ontology.

Some important factors used substantially in the research process are as follows:

– the opportunity offered by the Content Description Tool some sections (spatial parts) of the digital images to be annotated semantically besides the image as a whole;

– the specificity of the subject domain: cannons of creating Orthodox iconographical images, which give an explicit semantic connection between the character and her/his required/expected place on the particular image;

– OBIO ontology as a research experiment by its own, with open opportunity to modify its life-cycles.

The main idea behind the dynamic support of the annotation process is to provide the software a technique of "awareness" about the currently filled annotation and its similarity to the offered knowledge patterns. The process is interactive and on each step "best suggestions" are given to the annotator.

For example, if an image of "Deisis with the Apostles" icon is annotated the *Annotation View to Icon* will be used (see the Pattern Graph on Fig. 5). While the information of global annotation is filled, the image will be described as *compound image*. This will bring to the annotator the patterns presented in the ontology for compound images and particularly for "Deisis with the Apostles". The annotator will be provided with the formalized knowledge that the central picture presents an image of Christ, that on the left and on the right two columns with Apostles' portraits are placed. Each of the Apostles' portraits has a particular position among the others. If the annotator initiates a process of partial (segments) annotation, it is expected these segments to be attached to the known segments (the central image or a particular Apostle's portrait). Having the explicit knowledge of Apostles portraits positions and analyzing the incoming segments' spatial coordinates the Content

Description Tool should be able to suggest an initial knowledge pattern to the annotator, and probably even to "recognize" the portrait of an Apostle.

The ontology OBIO has to be extended with spatial semantic knowledge needed to support the interactive annotation process. Semantic notion of *Segment* should be introduced with a topological sense; also relations like **right of, left of, above, below,** etc., connecting the *Segments* in topological meaning; rules, to support reasoning in this sense and so on [7].

5. Conclusions

Large and ambitious projects as LOGOS require by researchers and knowledge engineers to construct knowledge models with pragmatic value.. Ontologies are promising constructions in knowledge sensitive research areas. Different formalisms for their implementation are used currently and the comparison among them тепърва ще се изследват предимствата на всеки от тях. The Content Description Tool of LOGOS Authoring Studio applies the Conceptual graphs formalism for annotating audio-visual content, combining both the practical implementation of annotation technology and research platform. During the final phase of the LOGOS project some enhancements of the ontology OBIO were experimented, targeting also applications beyond the LOGOS project.

References

1. A r a p i, P., N. M o u m o u t z i s, M. M y l o n a k i s, S. C h r i s t o d o u l a k i s. A Framework and an Architecture for Supporting Interoperability Between Digital Libraries and e-Learning Applications. – In: Proc. of First International DELOS Conference "Digital Libraries: Research and Development". Pisa, Italy, February 13-14, 2007, LNCS 4877, C. Thanos, F. Borri, L. Candela (Eds.). Springer 2007, 137-146.
2. C h e i n, M., M. -L. M u g n i e r. Conceptual Graphs: Fundamental Notions. – Revue d'Intelligence Artificielle, **6**, 1992, No 4, 365-406.
3. N. Crofts, Martin Doerr, Tony Gill, Stephen Stead, Matthew Stiff (Eds.). Definition of the CIDOC Conceptual Reference Model, March 2008.
   http://cidoc.ics.forth.gr/
4. D u B o i s, B. Towards an Ontology of Factors Influencing Reverse Engineering. – In: STEP'05: Proc. of the 13th IEEE International Workshop on Software Technology and Engineering Practice, USA, 2005, 74-80.
5. P a v l o v a - D r a g a n o v a, L., V. G e o r g i e v, L. D r a g a n o v. Virtual Encyclopeadia of Bulgarian Iconography. – International Journal "Information Technologies and Knowledge", Vol. **1**, 2007, No 3, 267-271.
6. S c h r e i b e r, G., H. A k k e r m a n s, A. A n j e w i e r d e n, R. D e H o o g, N. S h a d b o l t, W. V a n De V e l d e, B. W i e l i n g a. Knowledge Engineering and Management. – The CommonKADS Methodology, MIT Press, 2000.
7. S a r a y d a r o v a, V., K. S t a y k o v a, D. D o c h e v. Extensions of Logos Domain Ontology "Bulgarian Iconographical Objects". – In: Proc. of the Second LOGOS Open Workshop "Cross-Media and Personalised Learning Applications with Intelligent Content" LAIC 2008. P. Stockinger, D. Dochev (Eds.). 3 September 2008, Varna, 67-78.