

- [Pavlov&Paneva, '06] Pavlov R., Paneva D. (2006), Toward Ubiquitous Learning Application of Digital Libraries with Multimedia Content, *Cybernetics and Information Technologies*, vol. 6, № 3, pp. 51-62.
- [Pavlova-Draganova et al., '07] Pavlova-Draganova L., Georgiev V., Draganov L. (2007), Virtual Encyclopaedia of Bulgarian Iconography, *International Journal "Information Technologies and Knowledge"*, vol.1, №3, pp. 267-271.
- [Paneva et al., '05] Paneva, D., Pavlova-Draganova L., Draganov L. (2005), Digital Libraries for Presentation and Preservation of East-Christian Heritage, *Proceedings of the Second HUBUSKA Open Workshop "Generic Issues of Knowledge Technologies"*, Budapest, Hungary, pp. 75-83.
- [Pavlov et al., '06a] Pavlov R., Pavlova-Draganova L., Draganov L., Paneva D. (2006), e-Presentation of East-Christian Icon Art, *Proceedings of the Fourth HUBUSKA Open Workshop "Semantic Web and Knowledge Technologies Applications"*, Varna, Bulgaria, pp. 42-48.
- [Pavlov et al., '06b] Pavlov, R., Paneva D., Pavlova-Draganova L., Draganov L. (2006), Digital Libraries with Multimedia Content and Applications in Bulgarian Cultural Heritage (Analytical study), State Agency for Information Technologies and Communication (SAITC), by contract № 8/21.07.2005 between IMI-BAS and SAITC, 2006, Sofia, Bulgaria, Available at: http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf (accessed on April 10, 2007).
- [ICT, '07] Information and Communication Technologies, Work Programme 2007-08, Available online: ftp://ftp.cordis.lu/pub/fp7/ict/docs/ict-wp-2007-08_en.pdf (accessed on April 10, 2007)
- [Fensel, '04] Fensel D. (2004), *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*, Second edition.
- [Gruber, '93] Gruber T. (1993), Towards Principles of the Design of Ontologist Used for Knowledge Sharing. *International Journal of Human Computer Studies* 43 907-928, Available at: http://www.itee.uq.edu.au/~infs3101/_Readings/OntoEng.pdf (accessed on April 10, 2007).
- [Luchev, '06] Luchev, D. (2006), Experimental Digital Library - Bulgarian Ethnographic Treasury, *Proceedings of the Modern (e-) Learning Conference*, Varna, Bulgaria, pp. 106-111.
- [Luchev, '05] Luchev, D. (2005), Experimental Digital Archive "Bulgarian Ethnographic Treasury" (Analytical study) by contract № 12/08.09.2005 between SAITC and EIM-BAS, IMI-BAS (subcontractor), Sofia, Bulgaria, Available at: http://mdl.cc.bas.bg/DigLibr_BulEthnoTreasure_Analytical_Study_Luchev.pdf (accessed on April 10, 2007).

Authors' Information

Detelin Luchev - Ethnographic Institute with Museum, BAS, Bulgaria, 1000 Sofia, Moskovska str. 6A;
e-mail: luchev_detelin@abv.bg

Desislava Paneva - Institute of Mathematics and Informatics, BAS, Bulgaria, 8, Acad. G. Bonchev str.;
e-mail: dessi@cc.bas.bg

Konstantin Rangochev - Institute of Mathematics and Informatics, BAS, Bulgaria, 8, Acad. G. Bonchev str.;
e-mail: krangochev@yahoo.com

TOWARDS CONTENT-SENSITIVE ACCESS TO THE ARTEFACTS OF THE BULGARIAN ICONOGRAPHY

Desislava Paneva, Lilia Pavlova-Draganova, Lubomil Draganov

Abstract: *This paper presents an ontological model of the knowledge about Bulgarian iconographical artefacts. It also describes content-sensitive services for access, browse, search and group iconographical objects, based on the presented ontology that will be implemented in the multimedia digital library "Virtual encyclopedia of Bulgarian iconography".*

Keywords: *Ontology, Bulgarian Iconography, Digital Libraries, Content-sensitive services.*

ACM Classification Keywords: *I.2.4 Knowledge Representation Formalisms and Methods, H.3.7 Digital Libraries – Collection, Dissemination, System issues.*

Introduction

East-Christian icon art is recognised as one of the most significant areas of the art of painting. Regrettably, it is neglected in the digital documentation and the registry of the art of painting. This tendency is suspended by the team from the Institute of Mathematics and Informatics (Bulgarian Academy of Sciences) with the development of the multimedia digital library "Virtual encyclopedia of the Bulgarian iconography" (<http://mdl.cc.bas.bg/>). This valuable galleria of knowledge and specimens of East-Christian culture and art is created during the project "Digital libraries with multimedia content and its applications in Bulgarian cultural heritage" [Pavlov et al., '06b] [Pavlova-Draganova et al., '07] and includes several hundred specimens of Bulgarian icons and artefacts from different artists, historical periods, and schools [Pavlov&Paneva, '06] [Paneva et al., '05].

The impending development of the digital library points to the investigation and implementation of new techniques and methods for description of the semantics of iconographical artefacts and collections in order their valuable knowledge to be easy accessed and found. Knowledge technologies and Semantic web can provide these opportunities. Adopting them we make the first development steps of the "Bulgarian iconographical artefacts" ontology. This article presents the process of its consideration, scoping, and conceptualization, the ideas for its implementation in the "Virtual encyclopedia of the Bulgarian iconography". The interpretations of the iconographical knowledge do not have to be considered isolated from the standards and specifications in the field of cultural information representation. Therefore, section 2 summarizes the most important standard in the field of cultural heritage representation - CIDOC object-oriented Conceptual Reference Model (CIDOC CRM) and its use. Section 3 deals with different aspects of the ontology development, based on CIDOC CRM concepts and properties. Section 4 discusses content-sensitive services for access, browse, search and group of iconographical objects, based on the presented ontology that are in a process of implementation in the multimedia digital library "Virtual encyclopedia of Bulgarian iconography".

Ontological presentation of iconographical knowledge

One of the targets of the multimedia digital library "Virtual encyclopedia of the Bulgarian iconography" is to create rich context-based virtual presentation of the Bulgarian icon art and culture. Therefore, we observed and specified the experience that has been gained in the last 1000 years in the area of iconography to develop "a formal, explicit specification of a shared conceptualization" [Gruber, '93] about the iconography world - the ontology "Bulgarian iconographical artefacts". The annotator/indexers using this ontology will semantically describe and index the raw audiovisual iconographical content in order to create and maintain digital objects.

The interpretations of the iconographical knowledge do not have to be considered isolated from the standards and specifications in the field of cultural information representation because the goal is to maximize the reusability and portability of the designed ontological model. The most significant new development is the CIDOC Conceptual Reference Model, "object-oriented domain ontology" for expressing the implicit and explicit concepts in the documentation of cultural heritage. Since 9/12/2006 it is official standard ISO 21127:2006. It is the culmination of more than a decade of standards development work by the International Committee for Documentation of the International Council of Museums. Its role is to enable information exchange and integration between heterogeneous sources of cultural heritage information. CRM aims at providing the semantic definitions and clarifications needed to transform disparate, localised information sources into a coherent global resource. More specifically, it defines and is restricted to the underlying semantics of database schemata and document structures used in cultural heritage and museum documentation in terms of a formal ontology. It explains the logic of what they actually currently document, and thereby enables semantic interoperability. It intends to provide an optimal analysis of the intellectual structure of cultural documentation in logical terms.

The CRM is domain ontology in the sense used in knowledge technologies. It has been expressed as an object-oriented semantic model that can be readily converted to machine-readable formats such as RDF Schema, KIF, DAML + OIL, OWL, STEP, *etc.* It can also be implemented in any relational or object-oriented schema.

Real ontologies for concrete worlds of art objects are often developed as (conceptual at least) specializations of the CIDOC CRM ontology. During the creation of the "Bulgarian iconographical artefacts" ontology we observe the concepts and properties of CIDOC ontology and part of them we use in our ontology, other part we transform in order to fit for the iconography domain and several concepts don't belong to the CIDOC CRM ontology.

Concepts and properties in "Bulgarian iconography artefacts" ontology	CIDOC CRM chains
<p><i>Dimension</i></p> <p>Iconographical object → has dimension → Dimension Dimension → was observed in → Unit of Measurement Dimension → has width value → Number Dimension → has height value → Number Dimension → has length value → Number</p>	<p>E70 Thing (E22 Man-Made Object) → P43 has dimension (is dimension of) → E54 Dimension E54 Dimension → P40 was observed in → E16 Measurement (Unit of measurement of the dimension in our ontology) E54 Dimension → P90 has value → E60 Number (value of the dimensions of our ontology)</p>

Table 1: The concept Dimension and its properties in the "Bulgarian iconographical artefacts" ontology and the respective chain of CIDOC CRM concepts and properties

A juxtaposing example is shown in Table 1 for the concept 'Dimension' and its properties in the "Bulgarian iconographical artefacts" ontology and the respective chain of CIDOC CRM concepts and properties. The iconographical object can be adopted as a subset of the E22 Man-Made Object class. Ontology's concept 'dimension' is the same as CIDOC CRM E54 Dimension concept. The relationship between the iconographical object and its dimension is indicated by P43 has dimension property. In our ontology we adopted the 'Dimension→was observed in→Unit of Measurement' chain that is similar in CRM – 'E54→P40→E16'. In our ontology we split the 'E54→P90→E60' in three layers for the width, height and length of the iconographical objects [Pavlova-Draganova et al., '07].

Description of the semantics of the Bulgarian iconographical artefacts

Bulgarian iconographical domain contains a rich knowledge base that has to be semantically described. For this aim we observe and specify the experience in the area of iconography and start the development of the ontology presenting iconographical knowledge and artefacts.

The first activity in the process of the development of "Bulgarian iconographical artefacts" ontology is the definition of the scope of the ontology. Scoping has been mainly based on several brainstorming sessions with artists and content providers. It depends on the future implementation of the ontology in the multimedia digital libraries "Virtual encyclopedia of the Bulgarian iconography". These brainstorming sessions allowed the production of most of the potentially relevant terms. At this stage, we also juxtaposed these concepts to the available concepts in CIDOC CRM, thus concealing significant ambiguities and differences of opinion. A clear issue that arose during these sessions was the difficulty in discovering of definite number of concepts and relations between these concepts. The concepts listed during the brainstorming sessions were grouped in areas of work corresponding naturally arising sub-groups. Most of the important concepts and many terms were identified. The main work of building the ontology was then to produce accurate definitions.

The iconographical object is related to three levels of knowledge, enriched with a set of sub-levels of the data classification. All these levels of knowledge or "thematic entities" in the ontology conception are supported by the scientific diagnosis results and the related documentation [Pavlova-Draganova et al., '07].

- The entity "Identification" consists of general historical data, identifying aspects such as title, type, author, clan, iconographic school, period, dimensions, current location, description of the iconographical object/collection,
- The entity "Description" consists of information concerning the descriptive details of the theme and forms of representation, providing a better understanding of the context, such as characters and scenes, participation of characters in scenes, *etc.*
- The entity "Technical" includes technical information both revealing the techniques and the base materials used in the creation of the iconographical object/collection, and also concerning examinations of the condition, such as diagnosis or conservation treatments history.

These main entities and their metadata are supported, documented and provided by the scientific diagnosis, which has been applied to the iconographical objects and collections.

Figure 1 depicts the main classes and relations related to the concept 'Iconographical Object' in the ontology.

As it is shown on figure 1 in the "Bulgarian iconographical artefacts" ontology the concept 'Iconographical Object' is described with its title, author appellation, its clan and iconographic school, its current location and the period (time-span) of its creation, used base material and iconographic techniques, overall description. The ontology also captures the characters and scenes depicted on the iconographical object (icon, plastic iconographical object,

mural painting, iconostasis, iconographic element in Psalm-book, *etc.*) in order to be defined its compoundness. Figure 2 depicts the main subclasses of the 'Character' class.

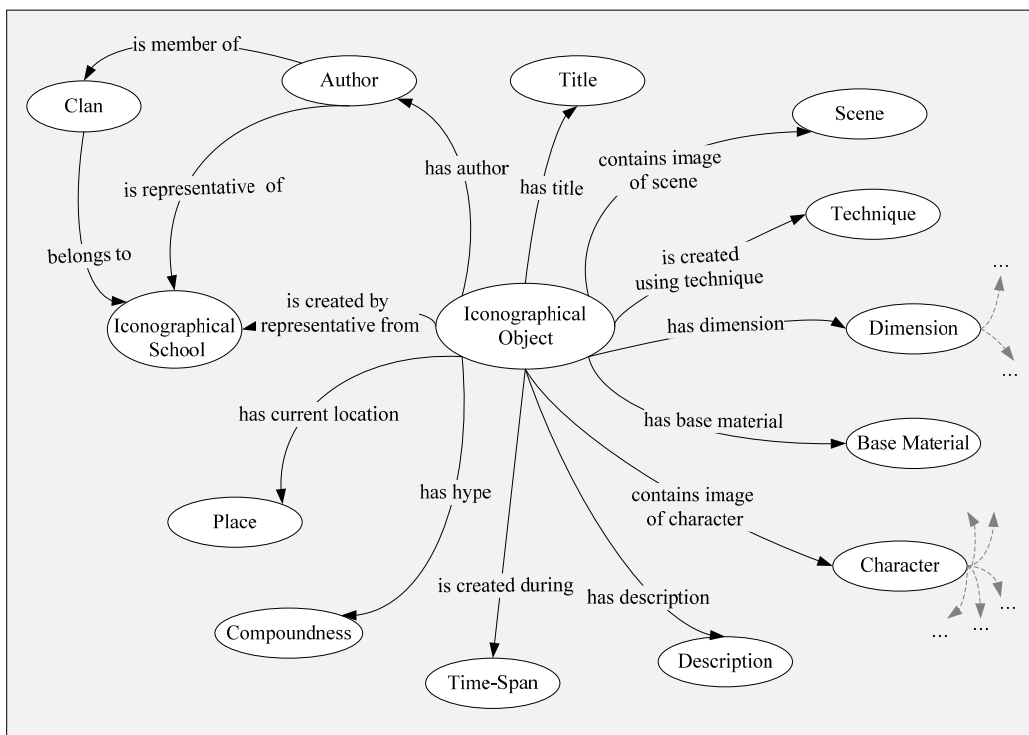


Figure 1: Main classes and relations related to the 'Iconographical Object' in the "Bulgarian iconographical artefacts" ontology

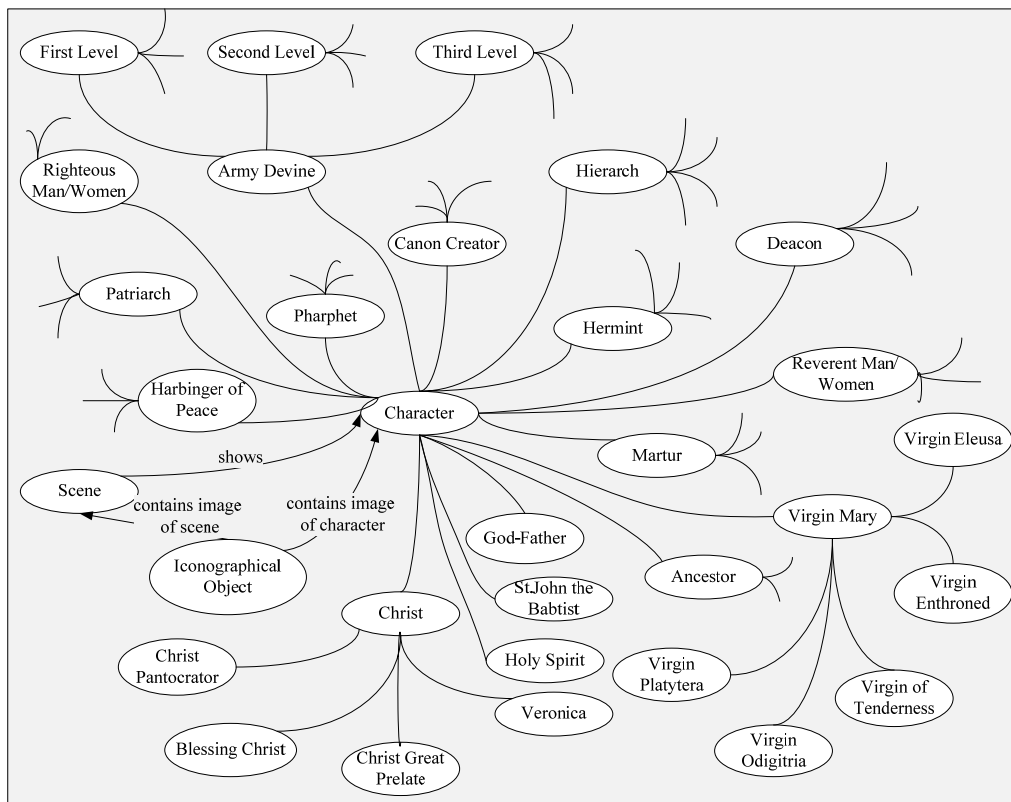


Figure 2: "Character" class and its subclasses in the "Bulgarian iconographical artefacts" ontology

New content-sensitive and customizing services in Virtual encyclopedia of the Bulgarian iconography

Multimedia digital library "Virtual encyclopedia of the Bulgarian iconography" currently provides its users with several services for present and search iconographical artefacts. But, the development of the "Bulgarian iconographical artefacts" ontology allows the inclusion of new semantic-based and content-sensitive access services with customizing elements in it.

One of them is "semantic-based search with grouping" depicted on a figure 3. It provides searching for iconographical artefacts that are created by representatives from chosen iconographic school, for example "Tryavna iconographic school". The results are lists of artefacts grouped according the several chosen criteria: authors, title, period, location, base material, depicted characters and scenes. This grouping opportunity will be very helpful for quick find of definite artefact in the iconographical object repository. During the search process the semantic-based service traces nodes of the ontological tree and presents instances of checked classes.

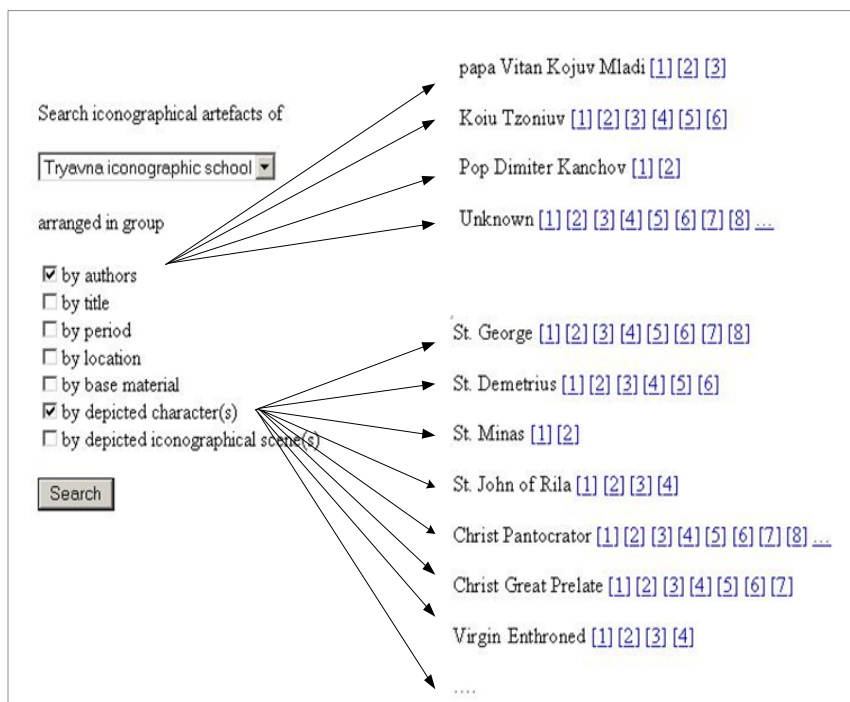


Figure 3: "Semantic-based search with grouping" in multimedia digital library "Virtual encyclopedia of the Bulgarian iconography"

Similar service could provide this grouping artefacts

functionality during the multi-criteria search (at present available in the digital library), but the desired grouping criteria have to be selected by the user during its personal profile creation. This action dictates the proper iconographical object observation style [Paneva, '06].

Another content-sensitive service is the "content browsing". It will display the ontology information graphically in order to support artists to easily navigate and browse through the concepts. Moreover, it provides them with information about the concepts and other related issues concerning the ontology. This service will be particularly useful to artists who are not familiar with concept searching and want to browse the information resources in a user-friendly way.

The displayed concepts will be obtained querying the ontology. If the artist requires the concrete content associated to any of the concepts displayed by the iconographical content browsing service, another query is done, this time, on the content database. In such way, the artist gets the information requested with the precise content to build his story.

Conclusion

The "Bulgarian iconographical artefacts" ontology tries to capture the knowledge in the iconography domain in order to provide tool for semantically description and indexing of the raw audiovisual iconographical content digital objects in the multimedia digital library "Virtual encyclopedia of the Bulgarian iconography". This ontology can be use for realization semantic-based access and search of concrete iconographical objects, as it shown in this paper. The future development of the digital library will continue to improve and extend the "Bulgarian iconographical artefacts" ontology and the DL services based on it.

Bibliography

- [Pavlov et al., '06b] Pavlov, R., Paneva D., Pavlova-Draganova L., Draganov L. (2006), Digital Libraries with Multimedia Content and Applications in Bulgarian Cultural Heritage (Analytical study), State Agency for Information Technologies and Communication (SAITC), by contract № 8/21.07.2005 between IMI-BAS and SAITC, 2006, Sofia, Bulgaria, Available at: http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf (accessed on April 10, 2007).
- [Pavlova-Draganova et al., '07] Pavlova-Draganova L., Georgiev V., Draganov L. (2007), Virtual Encyclopaedia of Bulgarian Iconography, International Journal "Information Technologies and Knowledge", vol.1, №3, pp. 267-271.
- [Pavlov&Paneva, '06] Pavlov R., Paneva D. (2006), Toward Ubiquitous Learning Application of Digital Libraries with Multimedia Content, Cybernetics and Information Technologies, vol. 6, № 3, pp. 51-62.
- [Paneva et al., '05] Paneva, D., Pavlova-Draganova L., Draganov L. (2005), Digital Libraries for Presentation and Preservation of East-Christian Heritage, Proceedings of the Second HUBUSKA Open Workshop "Generic Issues of Knowledge Technologies", Budapest, Hungary, pp. 75-83.
- [Paneva, '06] Paneva D. (2006), Use of Ontology-based Student Model in Semantic-oriented Access to the Knowledge in Digital Libraries, In proc. of HUBUSKA Fourth Open Workshop "Semantic Web and Knowledge Technologies Applications", Varna, Bulgaria, pp. 31-41
- [Pavlova-Draganova et al., '07] Pavlova-Draganova, L., D. Paneva, L. Draganov (2007), Knowledge Technologies for Description of the Semantics of the Bulgarian Iconographical Artefacts, In proc. of HUBUSKA Fifth Open Workshop "Knowledge Technologies and Applications", Kosice, Slovakia, 31 May – 1 June, 2007
- [Gruber, '93] Gruber T. (1993) Towards Principles of the Design of Ontologies Used for Knowledge Sharing. Int. Journal of Human Computer Studies 43, pp. 907-928, Available at: <http://www.itee.uq.edu.au/~infs3101/Readings/OntoEng.pdf> (Accessed on April 10, 2007).

Authors' Information

Desislava Paneva - Institute of Mathematics and Informatics, BAS, Bulgaria, 8, Acad. G. Bonchev str.;
e-mail: dessi@cc.bas.bg

Lilia Pavlova-Draganova – Laboratory of Telematics, BAS, Sofia, Bulgaria, 8, Acad. G. Bonchev str.;
e-mail: lilia@cc.bas.bg

Lubomil Draganov – Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria, 8, Acad. G. Bonchev str.;
e-mail: lubo@cc.bas.bg

AUTOMATIC CREATION OF LEXICAL RESOURCES FOR AN INTERLINGUA-BASED SYSTEM⁴

Juan Bekios, Igor Boguslavsky, Jesús Cardeñosa, Carolina Gallardo

Abstract: *The Universal Networking Language (UNL) is an interlingua designed to be the base of several natural language processing systems aiming to support multilinguality in internet. One of the main components of the language is the dictionary of Universal Words (UWs), which links the vocabularies of the different languages involved in the project. As any NLP system, coverage and accuracy in its lexical resources are crucial for the development of the system. In this paper, the authors describes how a large coverage UWs dictionary was automatically created, based on an existent and well known resource like the English WordNet. Other aspects like implementation details and the evaluation of the final UW set are also depicted.*

Keywords: *Lexical Resources, Wordnet.*

ACM Classification Keywords: *J.5. Arts and Humanities; H.2.8 Database Applications;*

⁴ This paper has been sponsored by the Ministry of Education and Science of Spain under project CICYT HUM2005-07260.