Table of Contents

International Journal of Computational Methods in Heritage Science

Volume 1 • Issue 1 • January-June-2017 • ISSN: 2473-5345 • eISSN: 2473-5337
An official publication of the Information Resources Management Association

Research Articles

The *International Journal of Computational Methods in Heritage Science* is indexed or listed in the following: Cabell's Directories; Google Scholar; Ulrich's Periodicals Directory

International Journal of Computational Methods in Heritage Science
Volume 1 • Issue 1 • January-June 2017

Retrieving Structured Information from (Semi-)/(Un-)Structured Cultural Object Documentation

Stella Markantonatou, Athena Research Centre, Athens, Greece
Panagiotis Minos, Athena Research Centre, Athens, Greece
George Pavlidis, Athena Research Centre, University Campus at Kimmeria, Xanthi, Greece

ABSTRACT

In the course of developing facilities for integrating cultural heritage in the everyday education practice, highly structured information was retrieved from both the structured and the unstructured Europeana documentation contributed by the Greek cultural institutions (~480K entries); Modern Greek is the working language. Satisfactory results were obtained by using in-house developed medium sized Getty/AAT compatible vocabularies and simple heuristics. The paper reports on the development of controlled vocabularies and the retrieval of structured information from the unstructured Europeana documentation. Retrieval results show the importance of controlled vocabularies and thesauri as regards the exploitation of digital library content.

KEYWORDS

Controlled Vocabularies, Digital Libraries, Education, Europeana, Multimedia Databases, Serious Games, Thesauri

INTRODUCTION

This work is about facilities for integrating cultural heritage into the everyday teaching practice; more particularly, it is about the development of platforms for creating serious games that take advantage of the cultural information in the web.

Two terms will be used throughout this paper: ‘learning object’ and ‘cultural object’ and they both denote how the respective digital objects have been documented rather than referring to their content. Thus, in the framework of this research, a learning object is a digital object retrieved from a repository that uses the international standards LOM[1]/LRE[2] to document the objects it contains. Similarly, a cultural object is a digital object retrieved from a repository that uses a CIDOC-CRM[3] (or some other international standard for cultural object documentation) compatible documentation. The underlying idea is that a learning object has been formulated and documented to address educational needs and it may contain one or more cultural objects, or even other learning objects for this matter of fact, that have not been developed or documented with education needs in mind (Markantonatou, Minos, Tzortzi & Pavlidis, 2016).

In addition to retrieval requirements, education software imposes quality control restrictions especially because it is interactive and open to the younger ages. A database that contains all the objects used by an educational system facilitates quality control, as opposed to free web access. Such a database has to ensure communication with international repositories of both cultural and learning objects;

DOI: 10.4018/IJCMHS.2017010106

therefore it has to be compatible with cultural object documentation standards such as CIDOC-CRM (Doer, 2003; Crofts, Doerr, Gill, Stead & Stiff, 2009) and learning object documentation standards, such as LRE-MAP[4]. A database designed in this way would accommodate both cultural objects and learning objects and would support the principled documentation and storage as well as the flexible search and retrieval of learning objects that contain cultural objects, for example serious games that contain pieces of music and 3D representations of statues and paintings. The authors of this paper have developed a database that satisfies these requirements and have populated it with cultural objects together with their standards compatible documentation that they retrieved from *Europeana*[5]. All the contributions to Europeana by Greek cultural institutions (>480K objects) were retrieved and stored.

This database required information that was more structured than the information available in Europeana. The quality of the structure of the information in Europeana varies with the provider. Quite often Europeana provides access to unstructured textual data – unstructured because they contain units of information that (1) from a standardization point of view, should have been codified under different rubrics (2) occur in unpredictable format and order – see the example cases (1)-(3) in the next section. Since structured information was required, the unstructured Europeana documentation had to be subject to some semantic analysis in order to make sure that the right information was accommodated in the right database slot. This is not the typical task of retrieving Europeana objects relevant to some description – such as the tasks discussed in (Petras, Ferro, Gäde et al., 2012; Petras, Bogers, Toms et al., 2013). Instead, for each object in Europeana, a new object was developed in the database; the new object had standardized metadata that were retrieved from both the standardized and the unstructured metadata of Europeana. The method applied drew on a combination of controlled vocabularies/thesauri and simple heuristics; satisfactory results were obtained.

It must be noted that neither conceptually organized lexica of considerable size and coverage, such as the WordNet[6], nor controlled vocabularies/thesauri of cultural terms nor authority lists of some considerable size are openly available for Modern Greek; therefore the resources used were developed from scratch. The Greek vocabulary of names of objects and materials that was developed was mapped on the Getty/AAT[7]; the mapping has assigned a structure to the Greek vocabulary and at the same time it constituted a first step towards a principled linking of Greek data with other Europeana data (de Boer, Isaac, Schreiber, van Ossenbruggen, Wielemaker & Stiller, 2011). AAT was chosen because it is a thesaurus for cultural documentation that covers a wide variety of objects, materials and techniques (Harping, 2010; Lanzi, 1998) and also, it has been used in multilingual data linking efforts within Europeana (de Boer et al., 2011).

In what follows, first the problem of retrieving structured information from the unstructured Europeana documentation is illustrated with examples. The retrieval method makes heavy use of controlled vocabularies and of mappings between vocabularies of different languages. As the different cultures have adopted somewhat different conceptualizations of the world, attempts to 'link' or 'map' vocabularies used to describe the cultural heritage of different societies run into correspondence problems; such is the case of mapping Greek controlled vocabularies extracted from cultural documentation texts with the AAT that more or less represents the American-English conceptualization of the world. Some indicative problems and the proposed solutions are discussed. Lastly, the method used to retrieve information from Europeana is presented and evaluated; the overall conclusion is that the proposed, relatively cheap yet resource-dependent, method of retrieval of structured information from the unstructured documentation of Europeana returns satisfactory results, at least in the case of the Greek data in Europeana.

STATEMENT OF THE PROBLEM

Consider the following Europeana entries:

1. http://www.europeana.eu/portal/record/2032008/207941.html

   a. **Title:** Το μνημείο του Λυσικράτη. Αθήνα, γύρω στα 1870 Πέτρος Μωραΐτης (ΦΑ_1_1261)
   b. **Provider's Translation:** The Lycicrates Monument. Athens, circa 1870 Petros Moraitis (ΦΑ_1_1261)
   c. **Type:** φωτογραφικό τύπωμα ("photographic print")
2. http://www.europeana.eu/portal/en/record/2032008/326280.html_
   a. **Title:** Νίκος Χατζηκυριάκος-Γκίκας (ΠΧΓ773) Σύνεργα του χτίστη στον κήπο, 1975 Μολύβι σε χαρτί, 0,22x 0,35 μ. Μουσείο Μπενάκη - Πινακοθήκη Γκίκα, Αθήνα
   b. **Provider's Translation:** Nikos Hadjikyriakos-Ghika (PHG773) Builders tools in the garden, 1975 Pencil on paper, 0.22 x 0.35 m Benaki Museum - Ghika Gallery, Athens
   c. **Type:** προσχέδιο ("preliminary drawing")
3. http://www.europeana.eu/portal/record/2032003/GR_item_64397.html_
   a. **Title:** {"gr"=>["Κεφάλι Αθηνάς"]} | {"gr"=>["Παρθένης Κωνσταντίνος"]}
   b. **Provider's Translation:** 'Athena Head', 'Parthenis Konstantinos'
   c. **Creator:** {"Gr"=>["Παρθένης Κωνσταντίνος"]}
   d. **Type:** {"gr"=>["ΣΧΕΔΙΟ"]}
   e. **Size:** Διαστάσεις έργου: 36 x 33,8 εκ. ("Dimensions")
   f. **Format:** Χρωματιστά κραγιόνια και τέμπερα ("Colored crayon and tempera")

The entries (1) and (2) have been provided to Europeana by the Benaki Museum[8] and the entry (3) by the National Gallery of Greece[9]. The English terms in double quote ("") are English equivalent terms retrieved from the AAT, they are not in the Europeana but they are provided here to facilitate the reader. Differences can be observed both across providers and in the material of the same provider. In all the entries the value of the field *Title* contains a substantial part of the information available about the object. This information, seen from the point of view of a cultural documentation standard such as the CIDOC-CRM, is a pastiche of the content of various semantic fields; the Europeana *dcTitle* definitely contains more than the actual title of a cultural object. So, in (1) the name of the cultural object is given first using a full stop as a delimiter and then the place is given, the date and the name of the artist. According to CIDOC-CRM, the last three pieces of information should not appear in the *Title* field but in three distinct designated fields. In (2) first appears the name of the artist with classification information as a delimiter, then the name of the cultural object with a comma as a delimiter, then the date, the information about the materials, the dimensions and the provenance of the object. Again, different types of information are packed in the *Title* field. In (3) both the title of the object and the creator are given in the value of the *Title* field. Still, (3) is more structured as regards the dimensions and the materials (*Format* field). The format of the dates has its own variations. Furthermore, different formats are used for proper names across providers; occasionally they conform to the format proposed by ULAN/Getty[10], that can be considered a near-standard for American English at least, but, what is important for our discussion here, proper name format is not consistent across the Europeana data. Last, in examples (1)-(3), the type of the object is consistently given as the value of the field *Type* but this is not always the case.

The task at hand, that would enable the database developed to support people involved in education to build learning objects using cultural objects, was to identify information about the title, the creator, the date, the materials, the type and the place (related with the object) and store it as values of the appropriate fields in the database. This information would be retrieved from the unstructured documentation of Europeana. To this end, simple rules were used, controlled vocabularies/thesauri and a mixture of heuristics.

CONTROLLED VOCABULARIES AND THESAURI

Three controlled vocabularies of Modern Greek were developed that provide support for the most frequent searches (de Boer et al., 2011): *names of cultural objects and of materials (COCD), creator*

names (CCD) and place names (PCD). The material in COCD was mapped on the Getty Art and Architecture Thesaurus (AAT). All the vocabularies provide preferred and equivalent terms. The result of the mapping of COCD on AAT is a new thesaurus that has the structure of the AAT as a backbone and encodes two different conceptualizations of the world, namely the Greek one and that of the AAT (probably the American English one).

Vocabulary of Names of Cultural Objects (COCD)

COCD was developed manually from museum documentation available on the web, including the Benaki Museum and the National Gallery of Greece. It contains 1000+ entries of which 900+ denote objects and 80+ denote materials (Table 1).

For the majority of the terms, it was possible to establish correspondences between the COCD entries and the AAT. The correspondences provided for multilingualism and for future activities in the domain of data linking. In addition, they provided for a structured version of the vocabularies. However, in the retrieval experiments only Greek and English terms and equivalent terms were used and no expansion.

As expected, the Greek culture has provided conceptualizations of the world and cultural terms that are not identical with the English ones; therefore COCD contains terms that cannot be mapped on the AAT; rather the AAT structure has to be minimally altered to accommodate those terms.

There were cases where the Greek language offered a less finely grained collection of terms for a semantic field. For instance, Greek seems to offer fewer terms for the collection of items that AAT classifies as narrower terms of bakeware: baking pans, bread pans, cake pans, springform pans, tube pans, cookie sheets, muffin pans, patty pans. Greek offers the term ταψιά for baking pans, φόρμες for all remaining 'pans' and the term λαμαρίνες for 'sheets'. In the solution adopted, bakeware has three immediate narrower terms, namely baking pans-ταψιά, φόρμες and λαμαρίνες. The remaining English narrower terms inherit from them according to their 'pan' or 'sheet' status respectively. A minor problem with this solution is that, in the new thesaurus, baking pans are placed higher than their former sisters (the remaining 'pans' and 'sheets'). Generally, however, at the level of terms that denote objects (as opposed to terms that denote categories of objects) there were few mismatches given the rich term collection of AAT and the fact that the direction of mapping was from COCD into AAT.

More interesting mismatches were observed with terms that describe categories of objects in Greek but have no equivalent in English, like in cases (4) and (5) that follow. The new categories that were introduced stand in a mother (4) or a daughter relation with AAT categories that have been proposed as the closest translation equivalents of the Greek term in various dictionaries or documents such as the official translations of EU texts. Of course, there may be no (even remotely) relevant English term (5).

4. **Greek term: singular: *ασημικό* (transcription: *asimiko*), plural: *ασημικά* (transcription: *asimika*).** The definition of *asimiko* in the established Dictionary of Standard Modern Greek (DSMG)[11] is: 'ασημικό το [asimikó] Ο38 (συνήθ. πληθ.): σκεύος ή διακοσμητικό αντικείμενο από ασήμι'. In the above definition, DSMG clarifies that the term is usually in the plural and

Table 1. Size and development cost of vocabularies

Man-months	Number of terms
12	Objects: 913 Materials: 82 Creators: 1196 Places: 1200 Total entries: 3391

adds the information 'cooking utensil or decorative object made of silver'. The term *asimika*, with this denotation, is found in the literature on cultural objects (Fraggedaki, 2010; Ballian, 2001). There are dictionaries that suggest silverware as a translational equivalent of *asimika* and even official translations use the term wrongly, for instance in legal texts of the European Union[12]. The term *silverware* of English has the following AAT definition 'Articles made of silver, especially tableware of silver used in serving and consuming food and drink' and the following WordNet definition: '(n) silverware (tableware made of silver or silver plate or pewter or stainless steel)'. Thus, the Greek term *asimika* has a wider denotation than the English term *silverware*. To map the Greek term on the AAT, a new node *asimika* was defined as a child of the Getty/AAT node metalwork. metalwork that has the following definition: 'Visual works that are the products of working any kind of metal, particularly metal objects of artistic merit.' *asimika* has the same definition as metalwork and only differs n the type of metal: 'Visual works that are the products of working silver or silver plated materials, particularly objects of artistic merit made of these materials.' *asimika* is defined as the mother of the node silverware. Thus, in the new thesaurus, a silver bowl is a new term that inherits from *asimika* and from the existing AAT term bowls (vessels). On the other hand, a silver spoon will be documented as silverware.

5. **Greek term: singular: *χάλκωμα* (transcription: *chalkoma*) plural: *χαλκώματα* (transcription: *chalkomata*).** DSMG provides the definition of *chalkoma* (singular) 'cooking utensil made of copper or an alloy of copper' and of *chalkomata* (plural): 'the set of cooking pots'. The term *chalkomata* is very rarely used to denote other objects than culinary equipment, for instance decorations made of copper (Chatzimichali, 2010). It is normally used to denote culinary equipment made of copper or alloys of copper in museum classifications, for instance, in the organization of the collections of Korgialeneio Museum[13]. Thus, for cultural documentation, the definition adopted is 'cooking utensil made of copper or an alloy of copper' and a new node 'chalcomata' was defined in the new thesaurus as a child of the Getty/AAT node *ware* that is described as follows: 'Refers to a class of manufactured articles, generally of utilitarian type, usually specified by its material, use, style name, or other characteristic, such as ceramic ware or kitchenware. In archaeology, it refers to distinct groups of pottery types classified according to characteristics of temper, hardness, type of paste, and surface treatment, rather than by shape or decoration, such as Kamares ware.' In the new thesaurus, a copper bowl is a new term that has *chalcomata* and the existing AAT term bowls *(*vessels*)* as wider terms.

Vocabulary of Names of Creators (CCD)

CCD contains ~1200 names and was developed manually from annotated data available on the web (creators are often annotated clearly as such despite the not-so-well-structured Europeana documentation). Preferred terms were adjusted to Getty/ULAN specifications where, according to the American English tradition, the surname is given first and the Christian name second and are separated with a comma. Greek uses the inverse word order. Names were transliterated according to the ELOT (Greek Standardisation Organisation)[14] specifications and the Greek term was given as an equivalent term (Table 2, example1). Other forms of the noun were listed as equivalent terms (Table 2, all examples). For instance, the creator *Χρύσα* is known with her Christian name in Greek therefore her surname was placed in brackets.

Vocabulary of Names of Places (PCD)

~1200 place names were semi-automatically retrieved by processing the documentation of the 483036 Greek entities in Europeana with automata that modeled the following simple phrasal patterns 1-3:

1. άποψη + (άρθρο) + D
   'view' + (determiner)+D

Table 2. Creator name codification

	Preferred term	Equivalent term
1	Γεωργίου, Κωστής	Κωστής Γεωργίου, Κ. Γεωργίου
2	Albertinelli, Mariotto	Albertinelli Mariotto
3	Βαρδέα, Χρύσα	Χρύσα (Βαρδέα)

2. άποψη + (Άρθρο) + λέξεις + (Άρθρο) + D
   'view' +(determiner)+words +(determiner)) + D
   {words=χώρας 'country', μονής 'monastery', περιοχής 'area', πόλης 'town', χωριού 'village', επαρχίας 'county', νησιού 'island', τειχών 'walls', τείχους 'wall', λίμνης 'lake', χερσονήσου 'penisnula', ναού 'temple/church', αρχαιολογικού χώρου 'archaeological site'
3. words + (determiner) + D
   {words = περιοχή 'area', πτέρυγα 'wing', πλευρά 'side'}
   (D: string consisting of one or two words at least one of which is introduced with a capital letter).
   Pattern 1 returned place names with precision 96%, Pattern 2 with precision 71% and Pattern 3 with precision 84%.

RETRIEVING INFORMATION FROM EUROPEANA

Europeana documentation was parsed and standardized information was retrieved and stored in the standards compatible database:

1. The fields *dcTitle* and *dcDescription* were parsed for extracting standardized (structured) information
2. Data were stored in the database. To this end, correspondences were established between Europeana and the database. Table 3 gives a detailed picture of these correspondences

Retrieval was language sensitive. Table 3 shows the type of matches between the Europeana fields and the standardized database. Direct matches did not require any NLP processing and values were simply copied from Europeana to the database provided that Europeana contained values for the particular fields. *dcdescription, dcTitle* and their translations, as it has been shown with cases (1), (2), (3), often contained various types of information that should be separated and encoded in distinct fields. This information was identified and copied with simple NLP. Minimal morphological augmentation of the terms in the Greek COCD and simple pattern matching (see the case (6) below) was used to retrieve and appropriately store the title, alternative titles if they existed, place, material and type of the object. For dates, a set of regular expressions has been used (6). The two bullets before the last bullet of (7) contain the names of the 12 months in the variations encountered in Europeana. Last, the sequence 'τέλη|αρχές|μέσα' is translated as 'end|beginning|middle' and 'Γύρω στα' is translated as 'about|circa'.

6. Regular expressions for retrieving dates

   - "(([|Γύρω στα) [0-9][0-9][0-9][0-9](-[0-9][0-9][0-9][0-9])?([πμ]\\.([ ]+)?X\\.)?)"
   - "([0-9][0-9][0-9][0-9](-[0-9][0-9][0-9][0-9])?([πμ]\\.([ ]+)?X\\.)?)"
   - "(((τέλη|αρχές|μέσα|Αρχές|Μέσα|Τέλη))?[0-9]+(ος|ου|ολού) αι\\.([πμ]\\.([ ]+)?X\\.)?)"
   - "([0-9]+(ος|ου|ολού)-((τέλη|αρχές|μέσα))?[0-9]+(ος|ου|ολού) αι\\.([πμ]\\.([ ]+)?X\\.)?)"

Table 3. Mapping EUROPEANA fields on the DB

DB field	Europeana object	Europeana field	NLP
object identifier	object	about	
object identifier (other)	object.proxy	dcIdentifier	
contributor	object.proxy	dcContributor	
creator	object.proxy	dcCreator	
object description (in Greek)	object.proxy	dcDescription	Y
object title (in Greek)	object.proxy	dcTitle	Y
object title translation	object.proxy	dcTitle	Y
Object title translation (other)	object.proxy	dctermsAlternative	Y
creation place	object.places	prefLabel	
creation place	object.places	altLabel	
creation place	object.places	hiddenLabel	
creation time	object.timespans	prefLabel	
creation time	object.timespans	altLabel	
creation time	object.timespans	hiddenLabel	

- "(((τέλη|αρχές|μέσα))?([0-9]+(ος|ου|ο)-((τέλη|αρχές|μέσα))?[0-9]+(ος|ου|ο) αι\\.([πμ]\\.([ ]+)?X\\.)?))"
- "([0-9]+(ος|ου|ο|ού)-[0-9]+(ος|ου|ο|ού) αι\\.)"
- "([0-9]+-[0-9]+(ος|ου|ο|ού) αι\\.)"
- "([0-9]+(ος|ου|ο|ού) αι\\.)"
- "([ABABαβab]' μισό [0-9]+(ος|ου|ο|ού) αι\\.)"
- "\\(([0-9]+-[0-9]+)([πμ]\\.([ ]+)?X\\.)?\\)"
- "([0-9]+/[0-9]+/[0-9]+)"
- "([0-9]+/[0-9]+/[0-9]+ - [0-9]+/[0-9]+/[0-9]+)"
- "((Ιανουάριος|Φεβρουάριος|Μάρτιος|Απρίλιος|Μάϊος|Μάιος|Μάης|Ιούνιος|Ιούλιος|Άυγουστος|Σεπτέμβριος|Οκτώβρης|Οκτώβριος|Νοέμβριος|Δεκέμβριος) (του)?[0-9]+)"
- "([0-9]+ (Ιανουαρίου|Φεβρουαρίου|Μαρτίου|Απριλίου|Μαΐου|Ιουνίου|Ιουλίου|Αυγούστου|Σεπτεμβρίου|Οκτωβρίου|Νοεμβρίου|Δεκεμβρίου) (του)?[0-9]+)"
- "([0-9]+ου π.Χ. αι.)"

EVALUATION

Two evaluation exercises of the NLP processing results were conducted (the correspondences that involved no NLP returned 100% success). In a first evaluation round, 100 entries of the DB that were retrieved from Europeana with NLP were collected by picking the first object entry in every 100 object entries and their contents were evaluated against the Europeana fields *dcDescription* and *dcTitle.* The values YES and NO were used as explained below.

1. YES: When material from Europeana *dcDescription* field fills the database field *Description* and the material from Europeana *dcTitle* fills the *Title* and *Alternative Title* fields of the DB.
2. YES: When Europeana fields *dcDescription* and *dcTitle* have no value and the database *Title* and *Alternative Title fields* have no value as well.

Table 4. Results of the first round of the evaluation

	Success	Failure (term missing)	Failure (other)
Type	68 (68%)	24 (24%)	8 (8%)
Place	91 (91%)	7 (7%)	2 (2%)
Material	100 (100%)	0	0

3. NO: When the material from Europena *dcDescription* does not fill in the *Description* field of the database and material from Europeana *dcTitle* does not fill appropriately the *Title* and *Alternative Title* fields of the database.
4. YES: When material from Euroepana *dcDescription* and *dcTitle* correctly fill the database fields: *type, material, date* and *place*
5. NO: When point 4 is performed wrongly.

The results of the first round of the evaluation are shown in Table 4.

The second round was conducted in the same way, this time for 484 entries of Europeana selected in the same way as in the first evaluation round and after the mapping algorithm had incorporated the heuristics. Table 5 shows the results of the second round of the evaluation.

Table 4 and Table 5 show that retrieval success is surprisingly stable across the two evaluation rounds. From this, it can be concluded that the Greek providers of Europeana have contributed relatively homogeneous material. Perhaps one reason for this homogeneity is that there are not many Greek contributors. To this fact, it must be added that Europeana texts use a particular sublanguage, that of cultural documentation. So, despite the fact that texts in *dcDescription* and *dcTitle* are unstructured from a standardization point of view, there is a regularity that facilitates retrieval.

CONCLUSION

In the framework of developing educational software that takes advantage of digitized cultural heritage, structured information was retrieved from the (semi-)/(un-)structured Europeana documentation. Documentation was considered 'structured' if it was compatible with documentation standards such as CIDOC-CRM and 'unstructured', if certain fields contained information that should be the content other fields.

Controlled vocabularies seem to be indispensable for this type of task. Good retrieval results were obtained from Europeana with small-medium sized vocabularies that require relatively small effort (the vocabularies of Modern Greek cultural object names, creators and places were developed from scratch on a 12 man-month effort) and rather simple pattern matching techniques. It is likely, though, that homogeneity of the Europeana contents and documentation, in particular the consistent use of a sublanguage, affected the results in a positive way.

Table 5. Results of the second round of the evaluation

	Success	Failure (term missing)	Failure (other)
Type	328 (67,7%)	130 (26,9%)	26 (5,4%)
Place	440 (90,9%)	34 (7%)	10 (2,1%)
Material	484	0	0

This said, the evaluation exercises have suggested that the main cause of failure was the size of the controlled vocabularies, so the authors' immediate plans include the enrichment of the population of entries and the improvement of the quality of information. A next step is the linking of the vocabularies developed by the authors with the Europeana pool of vocabularies. Lastly, it is planned that the vocabularies will be made publicly available, probably enhanced with a crowdsourcing system.

REFERENCES

Ballian, A. (2001). Αφιερωματικές επιγραφές και δωρητές στα εκκλησιαστικά ασημικά 17ου-19ου αιώνα. *ΜΟΥΣΕΙΟ ΜΠΕΝΑΚΗ 1*, 2001:87-110.

Chatzimichali, A. (2010). *Sarakatsani* (Vol. 2). Athens, Greece: Aggeliki Chatzimichali Foundation.

Crofts, N., Doerr, M., Gill, T., Stead, S., & Stiff, M. (2009). *Definition of the CIDOC Conceptual Reference Model: Version 5.0.1*. Retrieved from http://cidoc-crm.org/docs/cidoc_crm_version_5.0.1_Nov09.pdf

de Boer, V., Isaac, A., Schreiber, G., van Ossenbruggen, J., Wielemaker, J. & Stiller, J. (2011). *Europeana Connect*. D2.3.1 Multilingual mapping of schemes and vocabularies.

Doer, M. (2003). The CIDOC Conceptual Reference Module: An ontological Approach to Semantic Interoperability of Metadata. *AI Magazine*, *24*, 75–92.

Fraggedaki, M. (2010). *Cultural Management in Museums and Municipality Contribution. The case of "Technopolis of the Municipality of Athens"*. Athens, Greece: National School of Local Governments. Ministry of Interior, Decentralization and Electronic Administration.

Harping, P. (2010). *Introduction to Controlled Vocabularies*. Los Angeles, United States of America: Getty Publications.

Lanzi, E. (1998). *Introduction to Vocabularies. Enhancing Access to Cultural Heritage Information*. Los Angeles, United States of America: Getty Publications.

Markantonatou, S., Minos, P., Tzortzi, K., & Pavlidis, G. (2016). Metadata for Learning Objects containing cultural objects. *International Journal of Mediterranean Archaeology and Archaeometry*, *16*(5), 53–61.

Petras, V., Bogers, T., Toms, E., Hall, M., Savoy, J., Malak, P., . . . Masiero, I. (2013). Cultural heritage in CLEF (CHiC) (2013). In Forner P., Mller H., Paredes R., Rosso P., and Stein B. (Eds.), Information Access Evaluation. Multilinguality, Multimodality, and Visualization, LNCS (Vol. 8138, pp. 92-211). Berlin Heidelberg: Springer.

Petras, V., Ferro, N., Gäde, M., Isaac, A., Kleineberg, M., Masiero, I., . . . Stiller, J. (2012). Cultural Heritage in CLEF (CHiC) Overview 2012. *Proceedings CLEF-2012*.

Stiller, J., Petras, V., & Gäde, M. (2015). *Europeana Version 3. MS12: White Paper on Best Practices for Multilingual Access to Digital Libraries*. Retrieved from http://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_Version3/Milestones/Ev3%20MS12%20Multilingual_Access%20White%20Paper.pdf

ENDNOTES

1 The IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata, https://standards.ieee.org/findstds/standard/1484.12.1-2002.html
2 The Learning Resource Exchange, http://lreforschools.eun.org/web/guest/about
3 CIDOC object-oriented Conceptual Reference Model (CRM), http://www.cidoc-crm.org/
4 The Learning Resource Exchange - Metadata Application Profile, http://lreforschools.eun.org/web/guest/metadata
5 http://www.europeana.eu/
6 https://wordnet.princeton.edu/
7 The Getty Art & Architecture Thesaurus, http://www.getty.edu/research/tools/vocabularies/aat/
8 http://www.benaki.gr/
9 http://www.nationalgallery.gr
10 Union List of Artist Names Online, http://www.getty.edu/research/tools/vocabularies/ulan/
11 http://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides/index.html
12 http://www.linguee.com/greek-english/translation/αση μικά.html
13 Korgialeneio Museum, Cephalonia, Greece, http://www.kefaloniainfo.com/kefaloniaisland/mouseia/korgialeneio/index.html
14 ELOT (ΕΛΟΤ) 743 E2, Information and documentation - Conversion of Greek characters into Latin characters, 2001, https://sales.elot.gr/online/search/details.do?documentId=300010000020380

Stella Markantonatou is a Research Director with the Institute for Language and Speech Processing/Research Center "Athena". She has published extensively in the domain of Computational Linguistics (theoretical issues, Machine Translation, documentation of Cultural Heritage, Computational Lexica). Her current interests concern the study of MultiWord Expressions and Cultural Heritage Documentation.

Panagiotis Minos is a computer scientist. His interests include the development of parsing software and environment for electronic lexicography.

George Pavlidis received his PhD in Electrical Engineering, earning the distinction of the Ericsson Awards of Excellence in Telecommunications. He has been working for numerous R&D projects with emphasis on multimedia systems in culture and education. In 2002 he joined the 'Athena' Research Center, where he is now a research director, head of the Multimedia Research Group and head of research at 'Clepsydra' Cultural Heritage Digitization Center. His research interests involve 2D/3D imaging, CBIR, multimedia technologies, human-computer interaction, intelligent user interfaces, multi-sensory environments and ambient intelligence, 3D digitization and reconstruction, 3D-GIS and mixed/augmented/virtual reality. Dr. Pavlidis is a member of the Technical Chamber of Greece, of the Hellenic Researchers' Association, a senior member of the IEEE, and a founding member of the 'Athena' Research Center's Researchers' Association.

Call for Articles

International Journal of Computational Methods in Heritage Science

Volume 1 • Issue 1 • January-June 2017 • ISSN: 2473-5345 • eISSN: 2473-5337
An official publication of the Information Resources Management Association

MISSION

The **International Journal of Computational Methods in Heritage Science (IJCMHS)** aims to present innovative research on digital solutions and advancements in the cross-disciplinary field of cultural heritage research and documentation. Through a presentation of applied and theoretical research and case studies from around the world, IJCMHS intends to be a pivotal resource for anthropologists, archaeologists, museum curators, librarians, archivists, historians as well as academics and students studying topics relevant to cultural heritage and conservation.

COVERAGE AND MAJOR TOPICS

The topics of interest in this journal include, but are not limited to:

Cultural Informatics • Digital Archaeometry • Digital Reconstruction • Documentation • Heritage Diagnostics • Heritage Materials • Historic Metallurgy • Open Data for Cultural Heritage • Public Engagement • Sustainable Heritage • Virtual Museums and Exhibitions

ALL INQUIRIES REGARDING IJCMHS SHOULD BE DIRECTED TO THE ATTENTION OF:

George P. Pavlidis, Editor-in-Chief • IJCMHS@igi-global.com

ALL MANUSCRIPT SUBMISSIONS TO IJCMHS SHOULD BE SENT THROUGH THE ONLINE SUBMISSION SYSTEM:

http://www.igi-global.com/authorseditors/titlesubmission/newproject.aspx

IDEAS FOR SPECIAL THEME ISSUES MAY BE SUBMITTED TO THE EDITOR(S)-IN-CHIEF

PLEASE RECOMMEND THIS PUBLICATION TO YOUR LIBRARIAN

For a convenient easy-to-use library recommendation form, please visit:
http://www.igi-global.com/IJCMHS