The semantic enrichment strategy for types, chronologies and historical periods in searchculture.gr

Haris Georgiadis, Agathi Papanoti, Maria Paschou, Alexandra Roubani, Despina Hardouveli, Evi Sachini

National Documentation Centre / National Hellenic Research Foundation
Athens, Greece
{hgeorgiadis, apapano, mpasxo, arouba, dxardo, esachin}@ekt.gr

Abstract. Most aggregators face challenges regarding searchability, discoverability and visual presentation of their content due to metadata heterogeneity. We developed an innovative metadata enrichment and homogenization scheme that is both effective and user-friendly and we embedded it in the ingestion workflow of searchculture.gr, the cultural heritage aggregator of National Documentation Centre (EKT). Two key components of the enrichment scheme are semantics.gr, a platform for publishing vocabularies that contains a tool for massive semantic enrichment, and a parametric tool embedded in the aggregator for chronological normalization. We enriched and homogenized the aggregated content with respect to types and chronological information which subsequently allowed us to develop advanced multilingual search and browsing features, including hierarchical navigation on types and historical periods, searching and faceting on type, time span and historical period, a tag cloud of types and an interactive timeline.

Keywords: aggregator, semantic enrichment, linked data, automatic categorization, vocabularies, thesauri, cultural heritage, historical periods, time-driven search, temporal coverage, timeline

1   Introduction

National Documentation Centre (EKT) has created a cultural content aggregation infrastructure [1] that collects content from digital repositories and libraries and publishes it in a central public portal, *www.searchculture.gr*. Soon our aim became to increase searchability and discoverability by providing new means of search, filtering, browsing and visual presentation of the content, based on two key types of metadata, item type (dc:type) and chronological information (dc:date, dcterms:created or dcterms:temporal, depending on the collection). However, the original documentation is extremely heterogeneous across the collections making impossible to support these features without intensive semantic enrichment and homogenization.

The heterogeneity of dc:type derives from term variations (different languages, synonyms, mixing plural and singular etc) and from different documentation methodologies (ranging from extremely general terms, such as "exhibit", to very specialized terms, such as "oenochoe", a specific type of vase). When it comes to temporal fields,

some providers use *period labels* (as problematic as the aforementioned type labels), but most of them use *chronological values*, such as dates, centuries, years and their *interval* counterparts which are also highly heterogeneous due to the use of different time encodings, languages, time granularity or literal descriptors (e.g. "Early").

As a solution we developed a semantic enrichment and homogenization scheme. It is based on *www.semantics.gr*, a platform for publishing vocabularies and thesauri that includes *a mapping tool for content enrichment* and contextualization. The tool sets mapping rules from multiple metadata field values to vocabulary terms. It uses a self-improving automatic suggestion mechanism and additionally supports the curator when intervening in the enrichment process. We also extended our aggregator infrastructure with a *parametric tool for the normalization of chronological values* which is based on regular expression processing. As a result the content of searchculture.gr was enriched and homogenized in respect to types, chronologies and historical periods, allowing us to publish it as Linked Open Data and to enhance the portal with new multilingual search and navigation features.

2   Related Work

Many aggregators use semantic enrichment techniques to deal with heterogeneity. Europeana [3], aggregation platforms such as MoRe [2] , MINT [11] and LoCloud [12] and research projects such as PATHS [8] use automatic enrichment with terms of established vocabularies. Complete automated enrichment on structured fields (such as dc:type) adopts an "enrich-if-you-can" strategy, horizontally, resulting in low enrichment coverage and high percentage of mistakes [6]. Automated annotation methods on more descriptive fields (such as dc:title) also yield relatively poor results [8]. As a result, homogenization is not achieved and this is why aggregators usually avoid offering advanced ways of exploration (browsing, faceting), that would improve the discoverability and the visual presentation of the aggregated content. TMP tool [13] of the AthenaPlus project is a platform for creating vocabularies that offers a mapping functionality which allows users to define equivalent relations between concepts from different vocabularies. However, unlike semantics.gr, it supports only SKOS vocabularies (which is not very suitable for time periods) while the mapping tool cannot perform more complex mappings (such as mappings from multiple fields, mappings from keywords contained in descriptive fields, use of logical expressions for filtering) and does not have a self-improving auto-suggestion mechanism.

Particularly for temporal enrichment, some aggregators enrich items described with period labels using timespan vocabularies (e.g. [3]), suffering, though, the abovementioned problems. Some attempt to homogenize chronological values to some extend as far as they conform to specific date formats ([2][11]). In either case, unlike our enrichment scheme, they don't handle items with temporal information uniformly, i.e. they don't assign period labels to items described with chronologies (a complex and error-prone task as highlighted in [9][10]) and vice versa (chronologies to items described with period labels). Our enrichment scheme supports chronological search and browsing both by year ranges and historical periods across all items with chronological

metadata, either explicit (temporal fields) or implicit (e.g. keywords in titles), regardless whether they were originally described with chronologies or with period labels.

3   The semantic enrichment scheme used in searchculture.gr

Searchculture.gr is based on an aggregator infrastructure that contains a tool for harvesting metadata and digital files (OAI-PMH harvester), a system for validating content againsts specific requirements (validator), a platform that implements the ingestion workflow (aggregator) and the public portal [1]. The search engine of the public portal is based on the Apache Solr[1] indexing system. The internal data model used for metadata storage is compatible with the EDM schema [4]. The aggregator was recently enhanced with a semantic enrichment scheme. We enriched the aggregated content with terms from a vocabulary of cultural item types, with homogenized chronological values (years or year intervals) and with terms from a vocabulary of greek historical periods. Both vocabularies were created with specific assumptions to facilitate the enrichment process. Metadata records are enhanced with three separate *EKT fields*. Note that the original documentation is not modified and is normally indexed and searchable.

3.1   The mapping-based enrichment tool of semantics.gr

Semantics.gr was initially created by EKT as a platform where institutions can create and publish RDF-based vocabularies and thesauri of any kind (concepts, timespans, agents, places) or any schema (parametric schema definition support). The platform was enhanced with a mapping tool that allows aggregators to enrich their collections with vocabulary and thesaurus terms [5][7].

The tool has a GUI environment with advanced automated functionalities that help the *curator* easily define *Enrichment Mapping Rules (EMR)* per collection from distinct metadata values to vocabulary terms. The tool accesses collection metadata via OAI-PMH harvesting. After setting the EMR, they can be served on request via a REST API in json format. Subsequently, the EMR can be used by an aggregator to enrich the collection in a bulk and straightforward one-pass fashion.

The EMR are defined per distinct value of a predefined metadata field (for example dc:type or dcterms:temporal), which is called *primary field*. In special cases the curator can choose a second metadata field (for example dc:subject) to create more precise EMR in case the documentation of the primary field is poor. We call this metadata field *secondary field* and its values filters. For example, a metadata record may have a dc:type value "folklore object" but a dc:subject value "Jewel" that reveals a much more accurate type. The enrichment tool supports automatic suggestion of EMR which by default is based on string similarity matching between metadata field values and indexed labels of vocabulary entries (e.g. skos:prefLabel and skos:altLabel). The automatic mapping suggestion is very effective and efficient leveraging the indexing system of semantics.gr search engine, namely Apache Solr. The tool can be easily configured

---

[1] http://lucene.apache.org/solr/

to be loosely coupled to the aggregator search portal (using deep linking) allowing the curator to easily search the collection for items having the specific values on primary and secondary fields. The curator can create complex logical expressions on the filters of a vocabulary entry assignment in order to create finer and more precise EMR and avoid false positives. For instance, they can use the logical NOT operator for setting exceptions.

When the automatic suggestion function fails to produce correct rules, the curator can set mappings manually. The enrichment tool "remembers" manual assignments in order to improve the effectiveness of auto-suggestion in future.

In certain cases, the curator can choose a highly selective *descriptive* field (the number of its distinct values approaches the number of all items) as a secondary field, such as dc:title or dc:description, if the values contain words or phrases that can reveal the appropriate vocabulary entry. For example a dc:title "An amphora from Attica" implies that the item is a vase. The tool searches inside such values for specific words or phrases derived from the vocabulary terms and then exposes only the matches as filters (instead of the entire field values).

3.2 The enrichment strategy for item types

We enriched and homogenized the aggregated content of searchculture.gr using a hierarchical bilingual SKOS vocabulary of item types that we created and published in semantics.gr[2]. The vocabulary consists of 414 terms, many of them linked to Getty AAT[3] (via skos:exactMatch). Eventually, metadata records were enriched with a separate field *EKT type* that holds references to the vocabulary. The type enrichment of a collection involves the following actions: i) examination of the documentation quality of the collection to decide whether a secondary field is needed or not ii) registration of the repository in semantics.gr iii) creation of EMR for the collection iv) ingestion (or re-indexing) of the collection in the aggregator in order for the actual enrichment to take place. Depending on the collection, the enrichment is based on original values of "dc:type" and, for special cases, of "dc:subject" or "dc:title". Table 1 summarizes 3 different documentation qualities, namely Type-A, Type-B and Type-C, and the respective mapping methodology.

Table 1. Documentations classes & type enrichment methodologies.

Class	Documentation quality class description	Methodology
Type-A	Good documentation of dc:type.	**EMR**: primary field
Type-B	Insufficient documentation on dc:type for part or all the collection, useful dc:subject	**EMR**: primary and **secondary** fields
Type-C	Insufficient documentation on dc:type for part or all the collection , useful dc:title or dc:description	**EMR**: primary field and **descriptive secondary** field

---

We will demonstrate the mapping process with the following example. Suppose that an aggregator-institution wishes to enrich its collections with references to a SKOS vocabulary (*V*) previously published in semantics.gr. Vocabulary *V* contains the following 5 entries:

➔ **http://scs.gr/sculpture** **skos:prefLabel** "Sculpture"@en | "Γλυπτό"@el
    ➔ **http://scs.gr/figurine** **skos:prefLabel** "Figurine"@en | "Ειδώλιο"@el
➔ **http://scs.gr/Jewellery** **skos:prefLabel** "Jewellery"@en | "Κόσμημα"@el
➔ **http://scs.gr/vessel** **skos:prefLabel** "Vessel"@en | "Σκεύος"@el
    ➔ **http://scs.gr/vase** **skos:prefLabel** "Vase"@en | "Αγγείο"@el

For a *Type-A* collection, the curator initializes a new EMR form in the enrichment tool where he/she sets the metadata field dc:type as the primary field and choose *V* as the target vocabulary. Then, the enrichment tool harvests metadata records from the repository and creates a list of distinct dc:type values with their cardinalities (1st column of Table 2). Next, the curator triggers the auto-suggestion functionality which successfully maps 3 distinct dc:type values to the correct vocabulary entries. The curator assigns the correct vocabulary term for the remaining value manually. Finally, the curator confirms the EMR and the mapping phase is completed. In Table 2, label "auto" indicates that the EMR was automatically created.

Table 2. EMR for a Type-A collection

dc:type value	Entry from vocabulary *V1*	
sculpture art *(120 items)*	**http://scs.gr/sculpture**	**auto**
greek vases *(230 items)*	**http://scs.gr/vase**	**auto**
jewelleries *(135 items)*	**http://scs.gr/Jewellery**	**auto**
amphora *(100 items)*	**http://scs.gr/vase**	**manual**

Table 3. EMR for a Type-B collection

dc:type	Filters (dc:subject)	Entry from vocabulary *V1*	
ceramic objects *(101 items)*	amphora , vase, statuette …	**http://scs.gr/vase**	**auto**
		if filter in ["vase", "amphora"]	**auto**
		http://scs.gr/figurine	**auto**
		if filter in ["statuette"]	**auto**
exhibits *(55 items)*	earing, amphora, …	**http://scs.gr/Jewellery**	**auto**
		if filter in ["earing"]	**auto**
		http://scs.gr/vase	**auto**
		if filter in ["amphora"] & **NOT** in ["earing"]	**manual**

Type-B collection has insufficient documentation of the primary field (either for all or for some of the items) but has another metadata field (secondary) that can contribute in the enrichment process. An example is shown in Table 3. Focus on the first mapping rule for dc:type value "ceramic objects": a metadata record with this dc:type value will be enriched with the reference http://scs.gr/vase only if it has one of the following

dc:subject filters: "vase" or "amphora" or with the reference http://scs.gr/figurine if it has a dc:subject value "figurine". The auto-suggestion mechanism can easily set this EMR as long as there are vocabulary matches for these filters. Items with dc:type value "exhibits" will be enriched with http://scs.gr/vase if they have a dc:subject "amphora" but they do NOT have a dc:subject "earing" (suppose that an image shows an earring shaped as an amphora).

In a *Type-C* collection, the documentation of dc:type is very poor for some items, but its dc:title values may contain useful words or phrases. The enrichment tool will search all titles against a set of words derived from all the labels of *V* (e.g. skos:prefLabel and skos:altLabel) as well as the keywords from previous EMR assignments and will set only the matching words as available filters for each dc:type value. The rest of the mapping process is identical with the one described for *Type-B* collections.

3.3 The enrichment strategy for chronologies and historical periods

We enriched the aggregated content with homogenized (normalized) chronologies and with historical periods using a hierarchical bilingual vocabulary of Greek historical periods. Metadata records are enriched with two separated fields, *EKT chronology* and *EKT historical period*.

Table 4. Documentations classes & temporal coverage enrichment methodologies.

Class	Documentation quality class description	Methodology
Temp-A	Temporal field (dcterms:temporal) with period labels (e.g. "archaic era")	1) **EMR**, primary field (EKT historical period) 2) extract year span from voc term (EKT chronology)
Temp-B	Insufficient documentation on temporal field for part or all the collection, useful titles (e.g. "archaic vase", "sculpture from the hellenistic period")	1) **EMR**, primary field, descriptive secondary field (EKT historical period) 2) extract year span (EKT chronology)
Temp-C	Temporal field (dc:date, dcterms:created or dcterms:temporal) with chronologies (e.g. "1981", "late 12th c. AD", "1100-1200 AD")	1) **Normalization of chronology values** (EKT chronology) 2) Enrich with EKT historical period
Temp-D	Temporal field (dcterms:temporal) with some values containing historical periods and others containing chronologies	Items with chronological values are handled as Temp-C and the remaining as Temp-A or Temp-B

Depending on whether the original temporal documentation is based on period labels or chronologies, we adopted two fundamentally different enrichment strategies, *historical period-driven enrichment* and *chronology-driven enrichment*, respectively. The former involves setting EMR in the enrichment tool of semantics.gr, similarly to the enrichment procedure for types; items originally described with period labels, are mapped to vocabulary terms but now they are also enriched with the respective year

ranges. In the chronology-driven enrichment, chronological values are being homogenized into years or year ranges and then, based on the results, the items are enhanced with the corresponding terms from the historical periods vocabulary. The enrichment is based on the original values of a temporal field ("dc:date", "dc:created" or "dcterms:temporal", depending on the collection) and in special cases taking into account keywords in descriptive field values, such as of "dc:description" and "dc:title".

Table 4 summarizes 4 typical collection types, namely *Temp-A*, *Temp-B*, *Temp-C* and *Temp-D*, their qualitative characteristics with respect to temporal documentation and the enrichment methodologies used. The methodologies used for *Temp-A* and *Temp-B* fall into the historical period-driven enrichment strategy, while the methodology used for *Temp-C* falls into the chronology-driven one. *Temp-D* collections are handled using both strategies: items described with period labels are handled as *Temp-A* or *Temp-B* and items described with chronologies are handled as *Temp-C*. We use regular expressions to distinguish chronologies from period labels.

The vocabulary of Greek Historical Periods. We created a Greek historical periods vocabulary that ranges from 8,000 BC (Mesolithic Period) to present and we published it in semantics.gr[4]. It is hierarchical and bilingual (Greek and English) consisting of 94 distinctive terms. The schema of the vocabulary conforms to the edm:Timespan contextual class introduced by Europeana [4]. For each term, apart from the different labels (skos:prefLabel, skos:altLabel) the year range is also defined in properties edm:begin and edm:end.

We created the thesaurus taking into consideration reputable sources about Greek history as well as established vocabularies such as Getty AAT. Some periods have a strict local scope (e.g. minoan, cycladic and helladic periods) and as a result their year ranges tend to overlap. We call those periods *relative*. The rest of the periods cover the entirety of Hellenic territory and are less debatable with respect to their timespans. We call those *absolute*. In our vocabulary, absolute periods have neither overlaps nor gaps when they have the same parent and relative periods have at least one absolute ancestor.

Table 5. Enrichment steps for a Temp-A collection

dcterms:temporal	Step 1: EMR - primary field (EKT historical period)	Step 2: extract year span (EKT chronology)
Post-Byzantine Period	→ Ottoman Period	→1453/1821
Middle - Late Hellenistic Years	→ Middle Hellenistic Period - Late Hellenistic Period	→-220/-31

Historical period-driven enrichment. The aggregator enriches items originally described with period labels with the mapped historical periods from the vocabulary (step 1) and computes their chronologies according to the assigned periods (step 2). The temporal enrichment of a collection involves the following actions: i) registration of the repository in semantics.gr, ii) creation of the EMR for the time field of the collection

---

[4] http://www.semantics.gr/authorities/vocabularies/historical-periods/vocabulary-entries/tree

in the enrichment tool, as described in section 3.2 (only for the period label values; if there are chronological values as well, these are automatically ignored by the mapping tool using regular expression filtering) iii) ingestion (or re-indexing) of the collection in the aggregator in order for the actual enrichment to take place. Tables 5 and 6 illustrate examples of *Temp-A* and *Temp-B* collections. The result of each enrichment step is presented for each item.

Table 6. Enrichment steps for a Temp-B collection

dc:title, dc:description (secondary field)	Step 1: EMR -primary & descriptive secondary field (EKT period)	Step 2: extract year span (EKT chronology)
Archaic vase	→ Archaic Period	→ -700/-480
Hellenistic sculpture	→ Hellenistic Period	→ -323/-31

Chronology-driven enrichment. The effectiveness of chronology-driven enrichment is heavily based on the normalization of chronological values. We developed a parametric tool which is based on regular expression processing that can handle 4 classes of chronological patterns, namely, "century range", "century", "year range" and "year/date". An authorized user can create many regular expression patterns for each class in order to capture as many chronological formats as possible. A chronological pattern can include custom and predefined placeholders that are associated with lists of keywords in many languages thus eliminating the number of different patterns needed for each class. Predefined placeholders affect the actual normalization algorithm. For example the "early" placeholder, which applies to patterns of "century range" and "century" classes, may have custom keywords "early", "beginning of" and the Greek counterparts. We created 30 different chronological patterns[5] and we arranged them in a specific order, from the stricter to the most ambiguous. When a chronological value is to be normalized, it passes through all the chronological patterns (or a subset of those that is set for the particular collection) sequentially, until the first match is found. Based on that pattern, the normalized year or year range is calculated. Table 7 shows some normalization examples per pattern class.

Table 7. Normalization of chronologies using regular expression matching

Chronological Pattern Class	Examples
century range	2nd half of 5th c. BC until 4th c. BC → **-450/-301**
century	early 18th century → **1700/1730**
	first half of 5th c. BC → **-500/-451**
year range	1342/48 → **1342/1348**
	1342 - 1654 → **1342/1654**
date/year	526 BC → **-526**

The aggregator normalizes the original chronologies (step 1) and based on the year or year ranges, chooses the corresponding *absolute* historical periods from the vocabulary

---

[5] Available at https://www.searchculture.gr/aggregator/resource/docs/Chronological_Patterns.pdf

(step 2). We did not use the *relative* historical periods, since we did not always have spatial information [9]. For example, Middle Bronze Age is an absolute period, covering the timespan 2000-1580 BC. It includes Middle Minoan, Middle Cycladic and Middle Helladic periods which are marked as relative since they refer to different civilisations that flourished in different territories. Therefore, an item dated in 1700 BC will be assigned with the Middle Bronze Age term.The chronological-driven enrichment is completely automated since there is no need for creating EMR. Table 8 illustrates an example of a *Temp-C* collection.

Table 8. Enrichment steps for a Temp-C collection

dc:date, dctemrs:created or dcterms:temporal	Step 1: Normalize chronologies (EKT chronology)	Step 2:Enrich with corresponding period (EKT historical period)
Late 5th century	→ 471/500	→ Early Byzantine Period
7th c. B.C-mid 6th c. BC	→ -700/-551	→ Early Archaic - Middle Archaic Period
03/11/1980	→ 1980	→ Regime change

3.4   EKT fields: encoding and indexing

The visual representation of EKT type field (as shown in the item page of searchculture.gr) consists of one or more types (e.g. "Figurine, Souvenir"). However, in order for the search engine to support advanced hierarchical searching and faceting on types, we index super (more general) types as well using a separate auxiliary Solr field.

For the EKT chronology field, we used the Date Range Field of Apache Solr which supports time interval indexing, time range queries and interval facets[6]. Regarding the EKT historical period field, its visual representation consists of either one (e.g. "Hellenistic Period") or two – in case of period intervals – historical periods (e.g. "Middle Archaic Period – Late Hellenistic Period"). In order for the search engine to support advanced hierarchical searching and faceting on periods, we index implying periods (e.g. those between the upper and lower bounds of a period interval) as well as both super (ancestor) and sub (descendant) periods in separate auxiliary Solr fields.

4   Enriching the content of searchculture.gr – The results

More than 150K items of searchculture.gr - 98% of the content - were classified into a compact and balanced set of 130 types. Table 9 illustrates the number of collections and enriched items per documentation class (Sec. 3.2).

After the type enrichment, searchculture.gr was enhanced with new multilingual search and browsing functionalities that improve discoverability including searching by type using a list of values, hierarchical navigation and faceting on types and an interactive tag cloud (Fig. 1). The enrichment improved remarkably the searchability of

---

[6] https://cwiki.apache.org/confluence/display/solr/Working+with+Dates

the content as illustrated in the experiment shown in Fig. 2(a) where we compared the number of search results returned by searchculture.gr for 6 search keys in Greek before and after the enrichment. Then we repeated the experiment, this time using the same search keys in English, as shown in Fig. 2(b). Since the majority of the items were documented only in Greek, the improvement was even more impressive.

Table 9. Collections and number of items per type documentation class

Documentation Class		# of collections	# of items
Type-A:	sufficient existing dc:type values	27	42583
Type-B:	insufficient dc:type values – useful dc:subject	24	60181
Type-C:	insufficient dc:type – resorting to dc:title values	4	55912
Total		**55**	**158676**

Fig. 1. Bilingual search, filtering (facets) and navigation (browsing) on types

Fig. 2. Type enrichment: improve in searchability of 6 types in Greek (a) and in English (b)

A total of 107,003 items of searchculture.gr - the 66% of the aggregated content - were enriched with normalized chronologies and assigned with historical periods. Note

that 63,234 items did not have any explicit temporal information, however, we managed to enrich 8,387 of them by identifying keywords in their titles or other descriptive fields. Table 10 illustrates the number of collections per documentation class as introduced in Sec. 3.3 and the total of enriched items per class.

Table 10. Collection and number of items per temporal documentation class

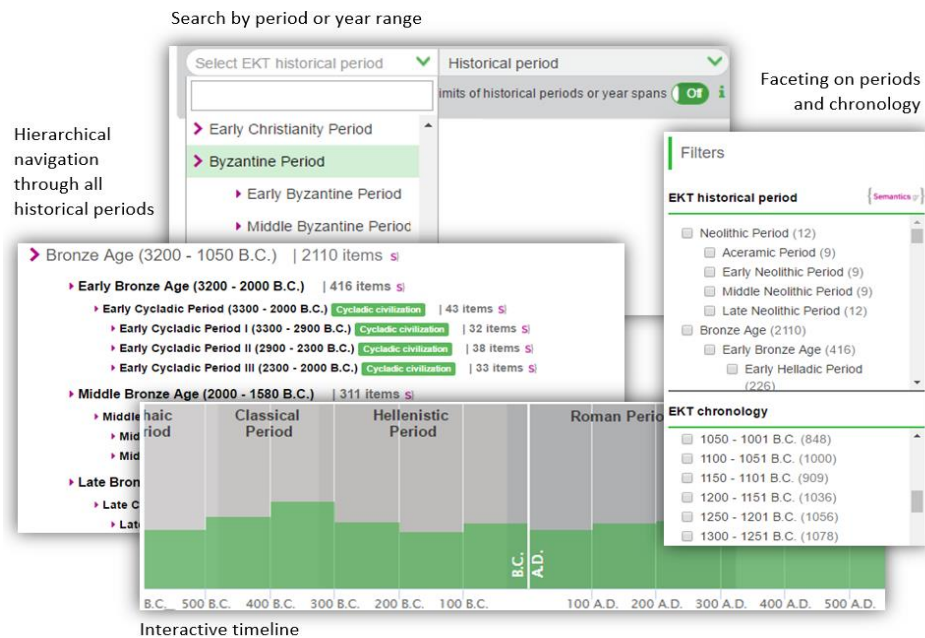| Documentation Class		# of collections	# of items
Temp-A	Temporal field (dcterms:temporal) with period labels	4	6870
Temp-B	Insufficient temporal field, useful titles or descriptions	3	6646
Temp-C	Temporal field with chronologies	44	81813
Temp-D	Mixed values: historical periods and chronologies	2	11674
Total		**53**	**107003**

Fig. 3. Bilingual search, filtering (facets) and navigation (browsing) on periods and chronology

After the temporal enrichment, searchculture.gr was enhanced with advanced time-driven multilingual search and browsing functionalities including searching by historical period and by year range, hierarchical navigation by historical periods, faceting on year-range and historical period and an interactive histogram-timeline (Fig.3).

Thanks to our EKT chronology and historical period indexing scheme (Sec. 3.4), users can choose between two modes for year or period interval search. In the "loose" one temporal search returns items with a year or period interval that overlaps that of the search criterion. For example, for a search criterion: "1500-1600 AD", an item dated "1550-1750 AD" will appear in the results. Similarly, for a search criterion "Classical

Period", an item dated "From Classical to Hellenistic period" will also appear in the results. In the "strict" mode, temporal search is more precise bringing only items with a year or period interval strictly within or coinciding the one defined by the search criterion. For example, for a search year range: "1500-1600 AD", an item dated "1550-1750 AD" will not be included in the results, while an item with "1550-1570 AD" date will. Similarly, for a search criterion "Classical Period", an item dated "From Classical to Hellenistic period" will not be included in the results, while an item dated "Early Classical Period" will. The "strict" mode is very useful when the user want to find items dated exclusively within as specific year or period interval.

Our future plans focus on extending the enrichment scheme in order to deal with spatial information and subjects. This will allow the enhancement of searchculture.gr with new features such as map-based navigation as well as searching, browsing and faceting on subject headings. Moreover, the multi-dimensional semantic enrichment will facilitate the creation of thematic exhibitions and "similar object" functionality.

References

[1] Georgiadis, H., Banos, V., Stathopoulou, I.O., Stathopoulos, P., Houssos, N., Sachini, E.: Ensuring the quality and interoperability of open cultural digital content: IISA: 178-183 (2014)

[2] Gavrilis, D. and Ioannides, M. and Theofanous, E.: Cultural Heritage Content Re-Use: An Aggregators's Point of View, ISPRS. II-5/W3: 83-87 (2015)

[3] Stiller, J., Petras, V., Gäde, M., Isaac, A.: Automatic Enrichments with Controlled Vocabularies in Europeana: Challenges and Consequences. EuroMed: 238-247 (2014)

[4] Europeana Data Model Primer, available at http://pro.europeana.eu/page/edm-documentation

[5] Georgiadis, H, Papanoti, A.:Semantics.gr: Information system for vocabularies & semantic enrichment. Technical report, EKT(2016)

[6] Manguinhas H. et al. Exploring Comparative Evaluation of Semantic Enrichment Tools for Cultural Heritage Metadata. TPDL 2016. Springer (2016)

[7] Georgiadis, H., Papanoti, A., et al: Semantics.gr: A self-improving service to repositories and aggregators for massively enriching their content. In Proc. of DHC Workshop of MTSR 2016

[8] Agirre, E., Barrena, A., Lopez de Lacalle, O., Soroa A., Fernando S., Stevenson, M.: Matching Cultural Heritage tems to Wikipedia. In: Proc. LREC 2012. Istanbul, Turkey. (2012)

[9] Rabinowitz, A.:It's about time: historical periodization and Linked Ancient World Data. ISAW Papers 7.22 (2014)

[10] Doerr, M., A. Kritsotaki , S. Stead: Which period is it? A methodology to create thesauri of historical periods. Proceedings of CAA2004 (2004)

[11] MINT. (2013). MINT Metadata Interoperability Services. http://mint.image.ece.ntua.gr/

[12] LoCloud enrichment, http://www.locloud.eu/Resources/LoCloud-enrichment-services

[13] Coudyzer, E. et al: The Terminology Management Platform: a Tool for Creating Linked Open Data. TOTh Workshop 2014