

Enriching the Greek National Cultural Aggregator with Key Figures in Greek History and Culture: Challenges, Methodology, Tools and Outputs

Haris Georgiadis^[0000-0003-1137-6583], Agathi Papanoti^[1111-2222-3333-4444],
Elena Lagoudi^[0000-0002-3431-4513], Georgia Angelaki^[0000-0001-6360-2664],
Nikos Vasilogamvris^[0000-0001-9121-2436], Alexia Panagopoulou^[0000-0003-3552-8932],
Evi Sachini

National Documentation Centre, Athens, Greece

Abstract. Since 2015, SearchCulture.gr, the Greek cross-domain Cultural Data Aggregator, a service developed by the National Documentation Centre in Greece (EKT), has collected a growing number of 800.000 digitised Cultural Heritage Objects (CHOs) from 73 cultural institutions. Addressing metadata heterogeneity in order to be able to provide advanced search, browsing and filtering options to users has been a key target from the start. Controlled linked data vocabularies for item types, historical periods and themes were developed over the course of the past years and are being used for the semantic enrichment of the CHOs' metadata. In the current paper we present the challenges, the methodology and tools used over the past 2 years for the process of enriching the aggregated CHOs' metadata with person entities from a Linked Data vocabulary comprising over 8.200 entries concerning Greek persons that left some mark in history, society, science, letters and art that we created for that purpose. This latest development allows all the works relating to a person to be interlinked and semantically enriched, adds significant browse and search functionalities to the portal and, therefore, opens new horizons for Greek SSH research.

Keywords: LOD Vocabularies, Digital Cultural Heritage, Data Repositories and Archives, Semantic Enrichment, Content Aggregators

1 Introduction

SearchCulture.gr (<https://www.searchculture.gr>) is the Greek National Aggregator for Cultural Data. It is being developed by the National Documentation Center of Greece (EKT), a public sector organization supervised by the Ministry of Digital Governance. Since its launch in 2015, it kept growing in numbers and expanding its functionalities. Today, SearchCulture.gr has amassed more than 800,000 records from 76 providers such as libraries, museums, archives and generally, any type of institution that is

custodian of cultural collections. Data ingested come from a variety of collections of archeology, folklore, history, arts and crafts and span more than 7,000 years of Greek history. SearchCulture.gr is the national accredited aggregator to Europeana.

With the goal to address metadata heterogeneity - the biggest challenge a large cross-domain aggregator faces – and in order to provide efficient search and browsing functionalities, SearchCulture.gr’s technical and operational strategy has relied heavily on semantic enrichments. We have developed an innovative metadata enrichment and homogenization scheme for types, subjects and temporal information and we embedded it in the ingestion workflow of SearchCulture.gr [1]. A key component of the enrichment scheme is Semantics.gr, a platform for publishing vocabularies as LOD. It also incorporates a Mapping Tool for enriching content en masse. We created and published three bilingual and hierarchical vocabularies for types, subjects and historical periods and we enrich and homogenize each collection we aggregate with terms from these vocabularies. This allowed us to develop advanced multilingual search and browsing features.

Our next step was to extend our enrichment scheme to persons, both creators and referred persons (i.e. a person depicted in a photograph, cast of a film, a recipient of a letter or the subject of a biography). Similar to the other enrichments, this would involve identifying person entities in metadata and mapping them to entries from a structured vocabulary, this time a catalogue of persons. The end goal was common to previous enrichments: to improve searchability and discoverability in SearchCulture.gr, by developing innovative person-driven search, browsing and faceting functionalities. This is important for the end user because with simple keyword-based search by personal names, the results are fuzzy: missing results because of the many different ways personal names are expressed (paronyms, given name diminutives, different languages, nicknames, the use of initials instead of full given names etc.) as well as false results that are related to synonymities. However, creating a vocabulary of persons and also, identifying, mining and disambiguating person entities in metadata was a much more complicated task which forced us to revise our enrichment scheme and strategy and to extend our tools in order to meet the new requirements.

2 Background: the enrichment scheme in SearchCulture.gr

The enrichment scheme in SearchCulture.gr is based on adding links stored in separate ‘EKT’ fields in *Cultural Heritage Objects*’ (CHOs’) metadata to terms from *Linked Open Data* (LOD) Vocabularies. These links are produced from explicit and curated mappings from distinct, aggregated metadata values to Vocabulary terms.

The scheme is based on Semantics.gr [2], a platform developed in-house by EKT for the development, curation and interlinking of vocabularies, thesauri, classification schemes and authority files- collectively hereby called *Vocabularies*- of any schema and their publication as LOD.

Semantics.gr also contains a *Mapping Tool* used to set *Enrichment Mapping Rules* (EMRs) in order to perform bulk data enrichment in aggregator databases and repositories. The GUI environment includes advanced automated functionalities that

help the curator easily define *EMRs* from source datasets (resources/terms from other vocabularies, metadata records or aggregated metadata values or phrases) to terms from a target vocabulary. For each dataset and target vocabulary a dedicated *Mapping Form* is created. The mapping form incorporates a self-improving automatic suggestion mechanism, though, a curator always validates the produced mappings. Moreover, the curator can correct, refine or create new mappings manually. In the context of SearchCulture.gr, the source dataset is always a set of aggregated metadata values or phrases, i.e. distinct metadata field values or distinct words/phrases contained in field values of a collection. These values are produced from metadata fields harvested via OAI-PMH. Validated mappings are served on request via a RESTful API in JSON format which can be used by the aggregator or repository to enrich the collection easily and en masse. The tool is thoroughly described in [1].

We started developing vocabularies and applying semantic enrichments to our collections in 2016. First, we enriched the original metadata using a Vocabulary of Item Types¹ that we created and published in Semantics.gr [1]. It is a SKOS-based LOD original vocabulary consisting of 193 terms that cover different types of cultural artifacts. Metadata records were enriched with a separate field "EKT type" that holds references to the vocabulary's terms.

The next step of the enrichment workflow was to homogenize and normalize chronologies and historical periods. The vocabulary Greek historical periods² was developed in 2017 and it is constructed according to the semantic class edm:Timespan of Europeana's EDM³. It contains 94 terms that cover Greek history from 8000 BC to today. It is hierarchical and bilingual. Original metadata records were enriched with two distinct fields, "EKT chronology" and "EKT historical period". Depending on whether the original temporal documentation was based on period labels or chronologies, we adopted two fundamentally different enrichment strategies [1].

The last iteration of enrichments was thematic, with a new field "EKT Subject" that includes references to terms of a bilingual and hierarchical vocabulary of subjects that is interlinked to the UNESCO Thesaurus⁴ (1,389 terms) and of a bilingual vocabulary of thematic tags⁵ that covers more specific topics (607 terms).

3 Adapting the enrichment scheme for person entities

The enrichments we conducted in the past and those of person entities vastly differ. The former were a type of classification using hierarchical vocabularies, allowing the curator some assumptions or generalization. The latter has the element of disambiguation and is much more demanding for the curator since each EMR potentially must be double checked and in some cases researched.

¹ <https://www.semantics.gr/authorities/vocabularies/ekt-item-types>

² <https://www.semantics.gr/authorities/vocabularies/historical-periods>

³ <https://pro.europeana.eu/page/edm-documentation>

⁴ <https://www.semantics.gr/authorities/vocabularies/ekt-unesco>

⁵ https://www.semantics.gr/authorities/vocabularies/thematic_tags

3.1 Two kinds of enrichments: creators and referred persons

The relation of a person with a CHO varies a lot. It could be the author or the subject of a book, the director, the screenwriter or the actor of a play, the sculptor, the model or the photographer of an antiquity, the sender, the receiver or the subject of a letter etc. The main problem is that many institutions do not clarify the role of a person and most importantly, even if they did in their repositories, most of them provide their metadata in aggregation schemata that are based on Dublin Core (such as EDM), which lack the required expressiveness to distinguish these roles. As a result, whichever is the specific relation of a person with a given CHO, the personal name will either appear in one of the two *agent fields*, “dc:creator” or “dc:contributor”, or it will appear in “dc:subject” or even in more descriptive fields such as “dc:title” and “dc:description”. Moreover, while the creator in most cases is given in “dc:creator”, persons that appear in photographs or are the subject of a book, sometimes appear in “dc:subject”, some other times in “dc:contributor” or “dc:title” or “dc:description”.

Ultimately, we decided that the enrichment on persons should improve SearchCulture.gr by allowing users to easily find all works of a creator (regardless of its specific role in the creation) and all the CHOs referring a person (regardless of the kind of reference). We opted for creating two separate fields, “EKT creator” and “EKT referred person” thus conducting those two kinds of person enrichments. These would be the basic fields that will be used for person-driven search, browsing and faceting.

3.2 The disambiguation problem

Disambiguation derived from synonymities is much harder when it comes to personal names. In the vast majority of our collections, references to persons are done with their names without any explicit identifying information (such as a link to an authority record) or explicit biographical information (such as the profession or the date of birth). It became quite obvious from the very beginning that implicit contextual information from other metadata (for example the type, the description or the date of a CHO) and sometimes the digital item itself (for example a portrait) should be taken into account by the curator in order to narrow down the options and make valid choices. Naturally, the vocabulary of persons that would be used should contain identifying and biographical information, not only for providing the end-user of SearchCulture.gr with useful information about a person, but mainly to be considered by the curator while making EMRs. The distinction between individuals with the same name is even harder when it comes to synonymities in distinguished political and economic families, where apart from the names, individuals may have common biographical information, such as a father’s name, occupation/profession or place of birth.

Another problem was the different forms of a personal name due to paronyms, given name diminutives, artistic nicknames (*noms de plume*), the use of initials instead of full given names, etc. Those issues made the following enrichment steps more difficult to be performed: i) string similarity-based comparison that is used for automatic mapping suggestions ii) mining of personal names from descriptive fields and iii) EMR curation. To solve these problems, on one hand we tried adding well-known paronyms,

diminutives and nicknames in person entries as alternative labels, and, on the other hand, we extended appropriately both the string similarity-based matching and the mining algorithms used in the Mapping Tool.

3.3 The scope

Given that SearchCulture.gr had already reached some 700K items at the beginning of this project, the number of personal names that appeared in the aggregated metadata was quite large. A decision needed to be made at the start as to the scope of the person enrichment project, for instance, whether it should include ancient creators or non-Greek important figures, etc.

For the first iteration of the enrichment project, it was decided to focus on Greek persons of importance, from antiquity to today. Some non-Greeks were included, because of their contribution and relevance to Greek history. It was also deemed right to focus on real people and not mythological figures and characters. Corporate bodies were also out of scope as were ensembles, artists' workshops etc.

It was also important to set some realistic targets about the coverage of the enrichments. In particular, we decided to focus mainly on persons who left some mark in Greek history, arts, science, politics and social life and whose biographical information would be reasonably easy to access. This was crucial not only for the time management of the project but also because a vocabulary without basic biographical information would prove useless in the disambiguation process.

3.4 Developing a vocabulary of Persons in Greek History and Culture

Unlike types and subjects, where the chosen level of granularity (how specialized the terms will be) can counterpoise the required number of terms in the respective vocabularies (since mappings can be done in broader concepts), a catalogue of personal names was expected to be much larger, if it was to achieve a broader enrichment coverage. Therefore, a big challenge was to come up with a vocabulary of persons with a critical mass of entries before the mapping processes started, although it was clear from the beginning that, as the mappings proceeded collection by collection, the vocabulary would be augmented with additional person entities.

One of the collections aggregated in SearchCulture.gr is "Pandektis: Modern Greek Visual Prosopography"⁶. It is a collection of more than 12,000 digitized portraits of Greek men and women who have attained distinction in different sphere of life, from the fall of Constantinople (1453) to the present day. Each record has metadata for the person's name (title and alternative title), place of birth and death, date of birth and death, and keywords that include information about the person's occupation or role in Hellenic history. Among those 12,000 records there are many duplicates, triplicates or more since the collection is not a catalogue of persons but a dataset of portraits and there can be multiple portraits of an individual.

⁶ <http://pandektis.ekt.gr/pandektis/handle/10442/14056>

After deduplication, we ended up with approximately 6,000 persons' records with basic metadata and (often multiple) thumbnails/portraits. These records became the core of our catalogue "Persons in Greek History and Culture"⁷, (from now on referred as *Vocabulary of Persons*). After evaluating the authorities that had been amassed and due to the nature of our content, we decided to also add in our records authorities of acclaimed art galleries' databases such as the National Gallery – Alexandros Soutsos Museum and the archive of Contemporary Greek Art Institute.

The Vocabulary of Persons is a LOD vocabulary conforming to the edm:Agent class of Europeana's EDM. It was created and published in Semantics.gr. Each entry was enriched when possible, with metadata regarding place of birth and death, date of birth and death, sex, occupation, bibliographic references and links to established resources, such as the Virtual International Authority File (VIAF), Wikipedia and IMDB Fig 1, illustrates an example of a person entity in view mode and as RDF.



Fig. 1. A person entity in the catalogue, and its RDF representation (in RDF/XML).

The catalogue initially consisted of approximately 6,500 entries. During the enrichment process we continued to add new entries. Today, it consists of 8,230 persons. The number keeps growing with every collection added in SearchCulture.gr.

In the Pandektis collection the field for occupation or role in Hellenic history is more or less homogenous. We decided to isolate and build on this information introducing a second vocabulary "Professions/Occupations"⁸. Use of this vocabulary facilitated various search and navigation functionalities. The "Professions/Occupations" is a LOD vocabulary conforming to the skos:Concept semantic class and consists of 366 terms. It is hierarchical and bilingual and its terms refer to occupations such as merchants, doctors and military officers, clergy positions, noble titles, different social movements affiliates like feminists or socialists or types of artists and literary creators. The terms of this vocabulary were used to classify entries in the Vocabulary of Persons.

3.5 The strategy for enriching aggregated collections en masse

As mentioned in Section 3.1, we conducted two kinds of enrichments, creators and referred persons. For each collection, we inspected the documentation particularities to

⁷ <https://www.semantics.gr/authorities/vocabularies/searchculture-persons>

⁸ <https://www.semantics.gr/authorities/vocabularies/professions-occupations>

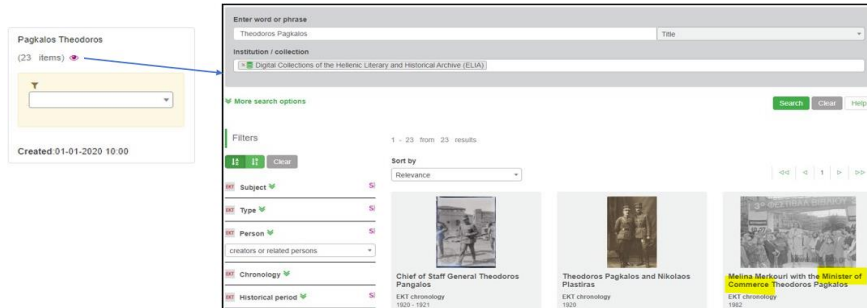
decide the metadata fields on which we would conduct mappings. For most collections, the creators' enrichments were done using one mapping form on the “dc:creator” field. However, for few collections creators’ enrichments were based on “dc:contributor”. Regarding the referred persons’ enrichments, we had to work on multiple mapping forms, usually for “dc:subject”, “dc:title” and “dc:contributor”. After we created all the mapping forms for every collection, we proceeded with the mapping processes.

Extending the matching algorithm used for automatic EMR suggestions. The mapping tool supports automatic suggestion of EMRs which by default is based on string similarity matching between metadata field values and indexed labels of vocabulary entries (preferable and alternative labels). The automatic mapping suggestion is based on a matching algorithm which is very effective and efficient leveraging the indexing system of semantics.gr search engine, namely Apache Solr. However, due to the different forms in which personal names are given (for example, the use of initials instead of full given names, use of patronymics – either in full form or as initials - between given names and surnames), it was necessary to adapt the matching algorithm in order to be as effective as possible. For example, when a personal name value includes initials, either for the given name or for the patronymic, these are used as patterns that match words that start with that initial. Moreover, since patronymics are usually optional and perhaps not included in the person entity record in the vocabulary, their existence in a personal name value should not filter our person entities lacking that patronymic. For example, values “Eleutherios Venizelos”, “E. Venizelos” and “Venizelos Eleftherios” will all match the personal entity “Eleutherios Venizelos”, but “Lefteris Venizelos”, “S. Venizelos” and “Sofoklis Venizelos” will not.

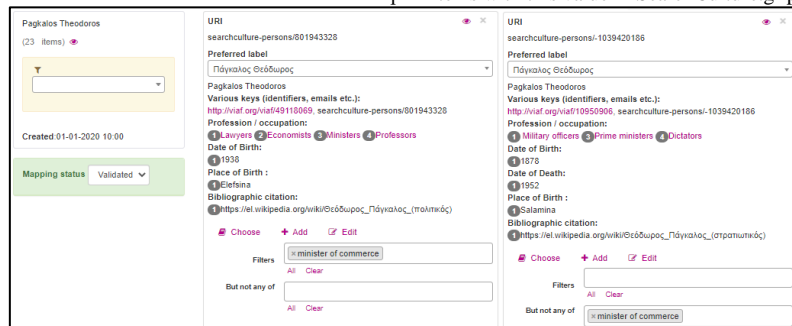
The screenshot displays a web-based mapping interface. On the left, a panel shows the original value 'Nop. Πωλ' and its mapping status 'Validated'. On the right, a detailed view of the selected person entity 'Nikolaïdis Nikos' is shown. This view includes a preferred label, an alternative label 'Nor Paul: Nop. Πωλ', various keys (identifiers, emails etc.), a list of professions/occupations (Lawyer, Publisher, Poet, Writer, Journalist), date of birth (1999), date of death (1991), place of birth (Athens), and a biographic citation. At the bottom, there are buttons for 'Choose', 'Add', and 'Edit'.

Fig. 2. An EMR set in creator mapping form.

Mapping forms on agent fields. Mapping forms in “dc:creator” and “dc:contributor” were easier to handle since their values are usually personal names. Fig 2, shows an EMR in a creator mapping form where the original value “Nop Πωλ” is mapped to person entity “Nikolaïdis Nikos”. Note that this EMR is automatically suggested since the “Nop Πωλ” (Nor Paul in English) is an alternative label in Greek.



(a) The curators click the “See items” link to inspect items with this value in SearchCulture.gr portal



(b) The final EMR assigns the 23 items to two person entities, according to the filter configuration

Fig. 3. An EMR on personal name value (found in title) refined by filters.

Mapping forms on non-agent fields. Field “dc:subject” is used for various kinds of thematic concepts including persons. More descriptive fields, such as “dc:title” and “dc:description”, are highly selective (the number of its distinct values approaches the number of all items). A mapping form for non-agent fields can be specially configured in order to search inside the values for specific words or phrases. The tool scans all distinct values of the field (e.g. all titles) and searches for inclusion matches against all the indexed labels of the Vocabulary of Persons. The search for inclusion matching is heavily based on the matching algorithm that is used for automatic EMR suggestions. Only the matching parts are considered as candidates for personal names and are exposed in the mapping form as values to be mapped. However, this technique will mine only personal names that match existing vocabulary entries. This is the reason why we first completed all mapping forms for “dc:creator” and “dc:contributor” for all collections, so as to add as many persons as possible in the vocabulary. Indeed, after all those mapping forms were processed, the vocabulary was enhanced by approx. 1,700 entries, which significantly increased the possibility that a personal name will be mined and given a chance to be mapped in non-agent fields in consecutive enrichments.

Due to the fact that the Vocabulary of Persons expanded in tandem to the enrichment process, it forced us to return often to several collections, re-harvesting new values in mapping forms and triggering new automatic suggestions in the hopes of mining new personal names and setting more EMRs.

Dealing with disambiguation. The mapping form provides tools that help curator in the disambiguation process. First of all, person entities from the vocabulary that appear in EMRs are presented in a panel which shows useful biographical information from the vocabulary record (see Fig. 2). The curator has immediate access to all this information that will help them in making the correct decision, without having to access the full person’s record.

Moreover, for each metadata value or phrase which is subject to mapping, there is a “see items” hyperlink that curators use to search SearchCulture.gr for items having the specific values in the metadata field in the specific collection. This way, the curator can easily preview all items that would be enriched according to the EMR (Fig. 3 (a)) [1].

If the curator discovers synonymities, they can use filters (values or phrases from the same or other metadata fields) and create logical expressions in order to fine-tune the EMR and avoid false positives [1]. For instance, they can use the logical NOT operator for setting exceptions. An example of this procedure is illustrated in Fig. 3 (b). The name “Theodoros Pagkalos” belongs to a military officer/ dictator (1878-1952) and to a politician (1938-). In order to disambiguate the two persons, we used as filter a phrase from the dc:title field. Only when a CHO has the phrase “Minister of Commerce” in the title it is enriched with the latter person.

3.6 Effort

We have enriched 77 collections of SearchCulture.gr with respect to person entities. For each collection, we created multiple mapping forms on various metadata fields, mainly “dc:creator”, “dc:contributor”, “dc:subject” and “dc:title” (167 forms in total). Table 1 shows the number of person names occurrences identified and mapped in original metadata as well as the number of validated EMRs per metadata field. The mappings took approximately 18 person months to complete, resulting in enriching a total of 177,986 CHOs out of the aggregated >836,000. 20% of SearchCulture.gr content was enriched, a considerable percentage given that the majority of our collections are archeological or folklore thus with few person references.

Table 1. Number of person names occurrences captured and number of validated EMRs

	Number of person names occurrences identified and mapped in original metadata	Number of validated EMRs
dc:creator / dc:contributor	229,156	10,325
dc:subject	39,633	982
dc:title /dc:description	59,946	5,845
Total	328,735	17,152

4 Person-driven search and browsing functionality

All our enrichments, are used as search and browsing criteria, thus facilitating access to the items via different pathways and producing targeted search results. The new

search box allows users to retrieve CHOs based on creators or referred persons, using a drop-down autocomplete list that presents entity name, a portrait and biographical information (Fig. 4 (a)). Another search field, using a drop-down autocomplete hierarchical list, allows users to retrieve CHOs created by or referring persons of a selected profession or occupation (Fig 4(b)).

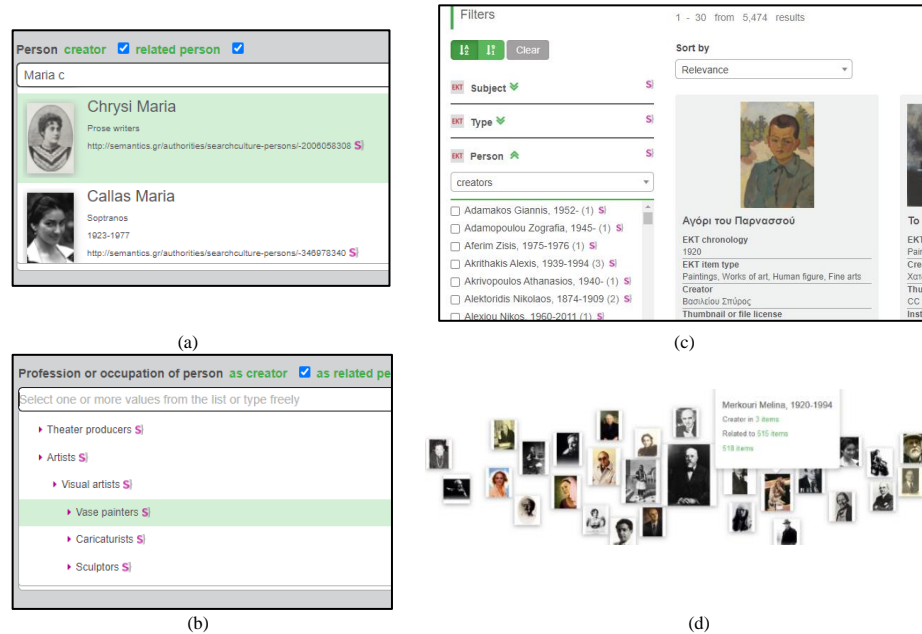


Fig. 4. New person-driven fields in the search-box (a, b), a person facet (c) and a tag cloud (d)

In the results' page, the enrichments are used as facets and serve to further narrow down the results. The newly-developed "persons" facet, showcases persons (either creators, referred persons or both) who appear in the results (Fig. 4 (c)). The facets can be used in combination, so one can search, for example, for "letters" exchanged between two persons by combining two different filters. On the homepage, an interactive navigation window displays a fraction of the persons as "tag cloud" of portraits (Fig. 4 (d)).

We also created a new search engine that targets the person entities in the aggregated content. The page is shown in Fig. 5. Browsing can be done by alphabetical order or by occurrence. Within this page one can also search for a name while autocomplete is enabled. In parallel, one can browse and select a value from the controlled list of occupations/professions and browse persons from that list. Alternatively, one can search by birth year or range of years, either in the search-box or in facets.

The resulting person entries include a thumbnail depicting the person, the name, key biographical information, the link to the person's entity in Semantics.gr and external links (VIAF, Wikipedia etc). On the right-hand side we see the number of CHOs the person has created, those that depict/refer to him/her and the total number of CHOs relative to this person. Numbers are clickable and direct to the relevant results' page.

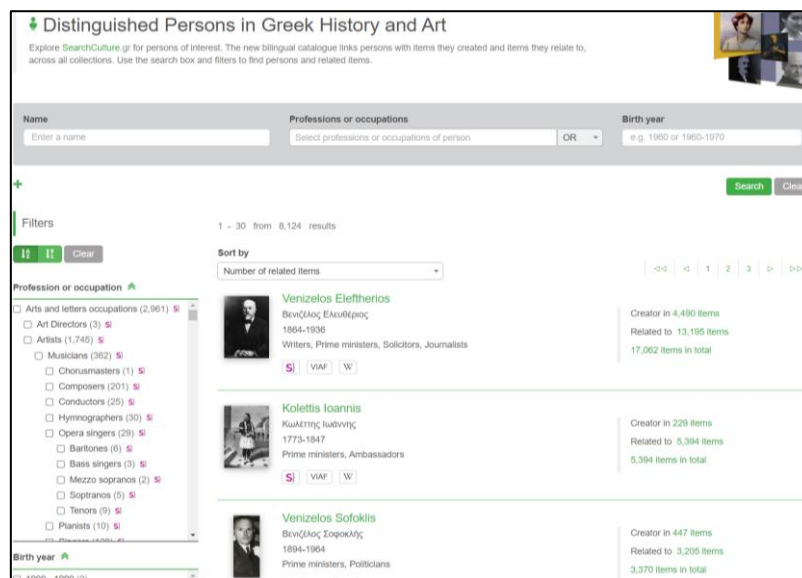


Fig. 5. A dedicated page for persons (<https://www.searchculture.gr/aggregator/persons/search>)

5 Related Work

Focusing in particular on the domain and thematic aggregators that form the Europeana Aggregators' Forum, some demand the data is enriched prior to ingestion, transferring the responsibility to the providers [3], others undertake semantic enrichment post-ingestion [7], while the majority just indexes string data without applying any semantic enrichment before delivering data to Europeana.

Europeana enriches aggregated CHOs by automatically linking text strings found in the metadata to controlled terms from LOD vocabularies [8][9]. Dbpedia is used for the enrichment of personal names in agent fields. Recently, Europeana curated an Entity Collection, a set of named entities harvested by controlled vocabularies, such as Dbpedia and Wikidata. Via this process source metadata are augmented with additional resources. This underpins some more controlled multilingual person-search functionality via the search-box.

Many EU-funded projects deal with the complexities of fully automatic or crowdsourced enrichments such as Enrich+⁹, St George on a Bike¹⁰, Europeana XX¹¹ and more. Other tools used in aggregation projects, such as MoRe [10] and MINT¹² cover mainly concepts, agents and places. However automated enrichment approaches

⁹ <https://pro.europeana.eu/project/enricheuropeana>

¹⁰ <https://saintgeorgeonabike.eu/>

¹¹ <https://pro.europeana.eu/project/europeana-xx>

¹² <http://mint-wordpress.image.ntua.gr/>

on structured fields adopt an “enrich-if-you-can” strategy, horizontally, resulting in low enrichment coverage and high percentage of mistakes [4] therefore unable to be exploited for building advanced search functionalities [11]. Our semantic enrichment scheme achieves high coverage and effective disambiguation because i) it can be adjusted to the documentation particularities of the individual collections ii) it combines self-improving automatic and fuzzy-based suggestions with a suit of tools that support the curation process and iii) uses a smaller and more compact target vocabulary.

Among other cross domain aggregators, the German Digital Library provides an advanced disambiguation and search functionality for persons, using the GND thesaurus as the source vocabulary¹³ and the OpenRefine tool for enriching person-values. Each person has a landing page with basic biographical information and the list of works of which one is the author or in which one is a referenced entity. Another worth mentioning tool developed by NTUA is SAGE¹⁴, which, similarly to the Mapping Tool of semenatics.gr, combines named entity recognition and string matching techniques with a human in the loop approach.

Finally, more targeted and narrow-focused in terms of metadata heterogeneity projects adopt a mixed higher curation strategy in which domain experts validate results provided with the use of AI tools [5][6].

6 Conclusions

The persons’ enrichment scheme along with the resulting search and browsing functionalities are quite unique for a large cross-domain cultural data aggregator as shown in the related work. Although the strategy and the methodology followed depended on human curation and was quite resource-demanding, the high quality and accuracy of the results justifies the effort.

Enabling person-driven search can be a crucial wayfinding strategy for widening access to the CHOs aggregated for communities such as teachers, students, researchers etc. Our ultimate goal is using contemporary semantic technologies in order to increase the content discoverability and to provide the means to extracting new leads, connections or insights.

Acknowledgments

The work presented in this article has been supported by the project “National Research Information and Technology System: Digital Content Aggregation, Documentation and Dissemination Infrastructure ensuring interoperability, long-term preservation and open access” of the Operational Programme “Reform of the Public Sector” (NSRF), co-funded by Greece and the European Union.

¹³ <https://www.youtube.com/watch?v=thK5ZjQbUR4>

¹⁴ <https://pro.europeana.eu/page/sage>

References

1. Georgiadis, H., Papanoti, A., Paschou, M., Roubani, A., Chardouveli, D., Sachini, E.: Using type and temporal semantic enrichment to boost content discoverability and multilingualism in the Greek cultural aggregator SearchCulture.gr. (2018) *Int. J. Metadata, Semantics and Ontologies*, Vol. 13, No. 1, pp.75–92
2. Haris Georgiadis, Georgia Angelaki, Elena Lagoudi, Nikos Vasilogamvrakis, Kostas Stamatis, Agathi Papanoti, Alexia Panagopoulou, Katerina Bartzi, Dimitris Charlaftis, Eleni Angelidi, Poly Karagianni, Despina Hardouveli, Evi Sachini: Publishing LOD Vocabularies in Any Schema with Semantics.gr. *MTSR 2021*: 280-291
3. Smith, M.: Linked open data and aggregation infrastructure in the cultural heritage sector, A case study of SOCH, a linked data aggregator for Swedish open cultural heritage, in *Information and Knowledge Organisation in Digital Humanities*, Routledge (2021), , <https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9781003131816-4/linked-open-data-aggregation-infrastructure-cultural-heritage-sector-marcus-smith>
4. Manguinhas H. et al. Exploring Comparative Evaluation of Semantic Enrichment Tools for Cultural Heritage Metadata. *TPDL 2016*. Springer (2016)
5. Koho, M., Ikkala E., Hyvönen E: Reassembling the Lives of Finnish Prisoners of the Second World War on the Semantic Web. *Proceedings of the Third Conference on Biographical Data in a Digital World (BD 2019)*, pp. 9 pp., *CEUR Workshop Proceedings*, 2022. bib pdf <https://seco.cs.aalto.fi/publications/2019/koho-et-al-reassembling-prisoner-biographies-2019.pdf>
6. Koho, M., Leskinen P., Hyvönen, E.: Integrating Historical Person Registers as Linked Open Data in the WarSampo Knowledge Graph. *Semantic Systems. In the Era of Knowledge Graphs. SEMANTiCS 2020* (Eva Blomqvist, Paul Groth, Victor de Boer, Tassilo Pellegrini, Mehwish Alam, Tobias Käfer, Peter Kieseberg, Sabrina Kirrane, Albert Meroño-Peñuela and Harshvardhan J. Pandit (eds.)), *Lecture Notes in Computer Science*, vol. 12378, pp. 118-126, Springer, Cham, Amsterdam, The Netherlands, October, 2020
7. Linked Open Data for Libraries, Archives and Museums, EuropeanaTech Insight, Issue 7, available at <https://pro.europeana.eu/page/issue-7-lodlam>
8. Manguinhas H., Isaac A., Charles V. et al: Europeana Semantic Enrichment Framework, technical documentation available at <https://docs.google.com/document/d/1JvjrWMTpMIH7WnuieNqcT0zpJAXUPo6x4uMBj1pEx0Y>
9. Stiller, J., Petras, V., Gäde, M., Isaac, A.: Automatic Enrichments with Controlled Vocabularies in Europeana: Challenges and Consequences. *EuroMed*: 238-247 (2014)
10. 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)
11. Agirre, E., Barrena, A., Lopez de Lacalle, O., Soroa A., Fernando S., Stevenson, M.: Matching Cultural Heritage terms to Wikipedia. In: *Proc. LREC 2012*. Istanbul, Turkey. (2012)