

Publishing LOD vocabularies in any schema with Semantics.gr

Haris Georgiadis¹, Georgia Angelaki¹, Elena Lagoudi¹, Kostas Stamatis¹, Agathi Papanoti¹, Alexia Panagopoulou¹, Katerina Bartzi¹, Nikos Vasilogamvrakis¹, Dimitris Charlaftis¹, Eleni Aggelidi¹, Poly Karagianni¹, Despina Hardouveli¹, Evi Sachini¹

¹ National Documentation Center (EKT),
Vasileos Konstantinou 48 Ave. 116 35, Athens, Greece

{hgeorgiadis, angelaki, elena.lagoudi, kstamatis, apapano, apanagopoulou, bartzi, nikvasil, dharlaftis, eage, pkarag, dxardo, esachin}@ekt.gr

Abstract. Semantics.gr is an innovative, cutting-edge platform developed by the National Documentation Centre (EKT) for the creation, curation, interlinking and publishing of vocabularies, thesauri, classifications and authority files (collectively referred to as *Vocabularies* in this paper) as Linked Open Data.

Besides SKOS, Semantics.gr can accommodate vocabularies of any data model that can be expressed as an OWL ontology. Through a user-friendly GUI, a user can represent a particular data model (i.e. SKOS or MADS/RDF) by configuring the relevant OWL classes and properties and setting appropriate constraints as a *Vocabulary Schema*. The Vocabulary Schema, in turn, provides the blueprint for the term forms through which a user adds or updates individual terms.

To date, we have represented various established data models in Semantics.gr, such as SKOS, EDM's contextual classes, MADS/RDF (for person names and organizations) and Schema.org, creating, in parallel, several custom Vocabulary Schemas capturing specialized vocabulary needs. We have also developed several Vocabularies that we use, on one hand, as controlled vocabulary lists in our digital repositories as well as to enrich 3rd party data in the national cultural aggregator, SearchCulture.gr, and the scientific data aggregator, OpenArchives.gr. Vocabularies are published as RDF and also via a Web API under an open license and are, therefore, free for any entity to reuse for their purposes in their own systems. Semantics.gr is launching as a free service to organizations that wish to create and publish their vocabularies as LOD, with the intention to become a central public platform for publishing and re-using authoritative LOD Vocabularies and to develop an active community around the service.

Keywords: LOD Vocabularies, Thesauri, Authority files, Linked data, Semantic enrichment, RDF, Vocabulary publishing service

1 Introduction

Semantics.gr is an innovative platform for the development, curation and interlinking of vocabularies, thesauri, classification schemes and authority files- collectively hereby called *Vocabularies*- and their publication as Linked Open Data (LOD). The infrastructure is developed in-house by the National Documentation Centre in Greece (EKT), alongside other services and tools, as per the organization's mission statement and its long-term strategy.

The platform has been in development since 2015 and was initially built with the scope to cater for in-house needs, i.e. to serve as a central infrastructure where EKT staff would develop the vocabularies that were being used to support the cataloguing and enrichment requirements of the organization's content repositories (such as the National PhD theses Archive, the EKT's Institutional Repository, the cultural and scientific data aggregators SearchCulture.gr and OpenArchives.gr, etc.) using semantic knowledge representation technologies.

The platform provides a smart, user-friendly interface that can be used even by non-experts in RDF, thus significantly lowering the barrier for institutions to publish vocabularies using semantic web technologies in a standardized way. The long-term goal for Semantics.gr is to become a central public platform for publishing LOD vocabularies, especially scientific terminology and authority files that can be further used by any third party in order to enhance the quality and the interoperability of their digital resources.

The platform is in production release available at <https://www.semantics.gr>. It currently includes all the functionality described in this paper as well as more than 40 published vocabularies freely available under open licenses for re-use. Some functionalities, such as the configuration of data models and schemata are currently available only to EKT personnel but the development and publishing of vocabularies will soon be also available, for free, to selected 3rd party legal entities (public, not-for-profit or commercial) that wish to develop and publish their own vocabularies as Linked Data. In the present article we examine the platform, its functionality and workflows for custom design of data models and vocabulary schemata and the development and publication of vocabularies. We also highlight some of the vocabularies developed by EKT scientific staff and how these are being used to homogenize and enrich our content repositories.

2 A Mechanism for Representing Data Models in a Customizable Way

Semantics.gr can support the publication of vocabularies that conform to any data model which can be expressed as an OWL ontology. This is achieved thanks to a user-friendly customization mechanism for the registration of data models and their specializations as concrete documentation schemata that adhere to the basic standards and specifications of the Semantic Web and LOD (RDF, OWL ontology, classes and properties). The mechanism is aimed to support the development of vocabularies that

are conformant either to the mainstream data models and ontologies or to customized data models that correspond to specialized documentation needs. A number of commonly used data models have already been integrated in the platform, such as SKOS², MADS/RDF³, the contextual classes of EDM⁴ and part of Schema.org⁵. At the same time, new data models have been developed such as oadm⁶ which was created by EKT as a model for the description of scientific resources and catriseu that was developed by the EU CatRIS project⁷ consortium as a model to describe research infrastructure services.

The data model chosen in each case generates the template on which a *Vocabulary Schema* is based to be created. The vocabulary schema in turn stipulates in detail the structure of the vocabulary that will be created in the infrastructure. In practice, the vocabulary schema defines the entry-update form through which the user adds or updates terms. These user-friendly forms are then used by authorized users assigned with the role of a *Vocabulary Curator* who creates and curates vocabularies. The forms possess control mechanisms that safeguard the constraints that are set for the individual properties by the chosen vocabulary schema. The forms can also be used for the curation of terms even by non-experts in RDF and OWL technologies.

Vocabulary curators can publish the vocabularies they have created as LOD following the adoption of a suitable open license. A published vocabulary is publicly searchable through the platform and is made available to third parties via the Web API and as a Linked Data.

2.1 Registering Data Models

Users with admin rights (currently EKT personnel) can represent any data model by specifying and customizing the OWL properties and then grouping them into OWL classes. A registered data model can then be used to produce the documentation schema that contains the exact structure and constraints underpinning the creation of one or more vocabularies.

In the user-friendly Administrator Environment the admin can register data models by configuring their namespaces, properties and classes. After the necessary namespaces (such as dc, dcterms and skos) are defined the admin can proceed to register properties by setting their namespaces, labels and constraints. For example, one can configure a particular property as referential, which means that it can be populated only with referential values (via their URIs) or with other nested resources (embedded resources, often without URIs which are depended on a parent resource) or even with language labels (via the xml:lang property in RDF/XML).

² <https://www.w3.org/2004/02/skos/>

³ <https://www.loc.gov/standards/mads/rdf/>

⁴ <https://pro.europeana.eu/page/edm-documentation>

⁵ <https://schema.org/>

⁶ OpenArchives Data Model (OADM), described here (in greek): <http://ariadne.ekt.gr/ariadne/handle/20.500.12776/16660>

⁷ Catalogue for Research Infrastructure Services (Catris): <https://www.portal.catris.eu/home>

Namespace	skos		
Name	broader		
Label	Broader (from this vocabulary)	English	×
	Ευρύτερη έννοια (από ίδιο λεξιλόγιο)	Greek	×
Comments	+		
	The identifier of a broader concept in the same thesaurus or controlled vocabulary. <code><skos:broader rdf:resource="http://www.mimo-db.eu/InstrumentsKeywords/4369_1"/></code>	English	×
Range category	Multiple resources		
Resource type	<input checked="" type="radio"/> URI Ref <input type="radio"/> Embedded resource		
Range for resources	<input type="radio"/> No constraints <input checked="" type="radio"/> Resources only of the same vocabulary <input type="radio"/> Resources only of the same contextual family <input type="radio"/> Resources only of the contextual family: <input type="radio"/> Resources only of the contextual classes: <input type="radio"/> Resources only of vocabulary: <input type="checkbox"/> Only authority semantic resources <input type="checkbox"/> Only deprecated semantic resources		
Inverse property	skos:narrower		

Fig. 1. The property editing form

Fig. 1 for example, depicts the `skos:broader` property registration form as defined by the SKOS data model. This property is referential, which means that it takes as value a URI that refers to another resource. An additional constraint is set in this particular case, which is that the resource referenced should pertain to the same vocabulary. Finally, as the `skos:broader` property is an expression of the bilateral relationship “generalization-specialization”, the inverted property of the relationship is also defined, that is `skos:narrower`.

Next, the admin can register OWL classes by setting their namespace and labels as well as the properties they contain.

Established classes and properties (under third-party namespaces), such as SKOS, are registered in Semantics.gr only with the purpose to support the creation and dissemination of vocabularies whose concrete schemata are based on these classes and properties. EKT personnel, who are granted with the permission to register data models and create schemata based on them, are responsible to respect the original semantics and specifications of the established data models and any possible configuration shall conform to that.

2.2 Building Vocabulary Schemata

After the classes and properties are registered, the user can configure a vocabulary schema which conveys the exact structure and additional constraints that will support the development of one or more separate vocabularies.

Initially, the user selects the basic class and any additional classes that the vocabulary terms will correspond to and decides which of their properties will be included in the schema. For every property, the user can select additional constraints. For example, they can narrow a referential property's value range to terms from a specific vocabulary published in Semantics.gr.

Fig. 2 shows the vocabulary schema entry/editing form that is based on the MADS/RDF model and particularly on the PersonalName class that is used to describe persons. In the schema quite a few properties take “nested” resources as their values. These nested resources are defined by other classes and their structure by separate schemata. They are incorporated in the description of the “parent” resource and it is not required to obtain URIs.

The screenshot displays a web-based form for configuring a property. The main form is titled 'madsrdf:identifiesRWO' and includes the following sections:

- Range category:** A dropdown menu set to 'Resource'.
- Resource type:** Radio buttons for 'URI Ref' and 'Embedded entity' (selected). Below is a text input field containing 'madsrdf:RWO'.
- Schema profile:** A dropdown menu set to 'madsrdf:RWO-PersonalName'.
- Index:** A dropdown menu and a checkbox labeled 'Index broader terms for terms of hierarchical vocabularies'.
- Property role:** A grid of checkboxes for roles such as 'Preferable label', 'Personal name', 'Corporate name', 'Alternative label', 'Same as', 'Identifier', 'Identifier scheme', 'Scheme type URI prefix', 'Scheme type ShortName', 'Earlier established Form', 'Later established Form', 'Begin', and 'End'.
- Hide:** A section with checkboxes for 'in form', 'in page view', 'in RDF', 'label from form', 'label from page', 'label from RDF', and 'Ref URI from RDF (only literal)'.

Below the main form, there are two nested schema details:

- madsrdf:hasIdentifier:** Schema for madsrdf:hasIdentifier. Base class: madsrdf:RWO. Range category: Multiple resources. Index: INDEX2. Identifier (checked).
- madsrdf:idScheme:** Schema for madsrdf:idScheme. Base class: madsrdf:Identifier. Range category: Literal or resource. Identifier scheme (checked).

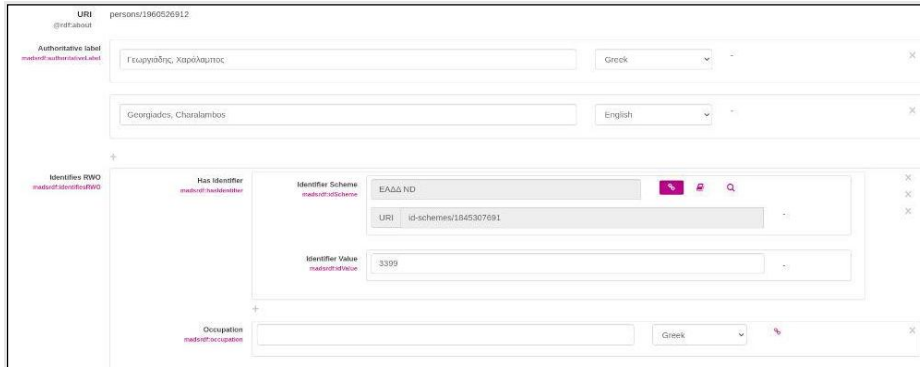
Fig. 2. Configuration of a property that incorporates a “nested” resource

Given the high level of complexity the functions described so far are accessible only to administrators from the EKT personnel who can represent and configure established or custom data models and vocabulary schemata and their adaptations in Semantics.gr. Upon request EKT admins can create new or adapt existing schemata to better serve any specific documentation needs of a third party that uses Semantics.gr to develop their own vocabularies.

The functions described in the following sections are accessible by vocabulary curators after they are granted the relevant permissions. Vocabulary curators can create and publish their own vocabularies using one of the pre-configured vocabulary schemata.

3 Creating and Publishing Vocabularies

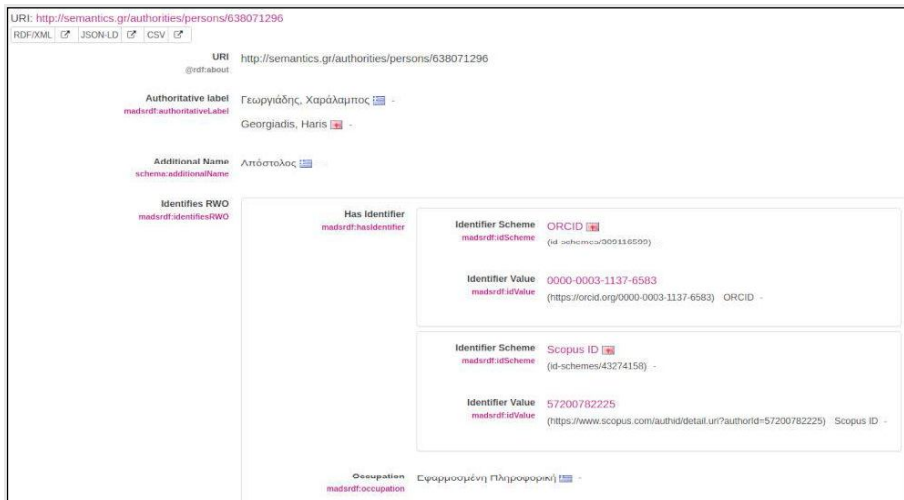
In order to start creating a new vocabulary, a vocabulary curator first selects the vocabulary schema that it pertains to and enters some descriptive properties such as creator(s), contributor(s), domain of application and license.



The screenshot shows the editing form for a `madsrdf:PersonalName` resource. The main form has the following fields:

- URI:** `persons/1966925912`
- Authoritative label:** Γεωργιάδης, Χαράλαμπος (Language: Greek)
- Additional Name:** Georgiades, Charalambos (Language: English)
- Identifies RWO:** A nested form with:
 - Has Identifier:**
 - Identifier Scheme:** EAD3 ND
 - Identifier Value:** 3399
 - Occupation:** Greek

Fig. 3. Editing form for a `madsrdf:PersonalName` resource



The screenshot shows the MADS/RDF resource page for `http://semantics.gr/authorities/persons/638071296`. The page displays the following properties:

- URI:** `http://semantics.gr/authorities/persons/638071296`
- Authoritative label:** Γεωργιάδης, Χαράλαμπος
- Additional Name:** Ληόσταλος
- Identifies RWO:**
 - Has Identifier:**
 - Identifier Scheme:** ORCID
 - Identifier Value:** 0000-0003-1137-6583
 - Has Identifier:**
 - Identifier Scheme:** Scopus ID
 - Identifier Value:** 57200782225
- Occupation:** Εφαρμοσμένη Πληροφορική

Fig. 4 A MADS/RDF `PersonalName` resource page

Then he/she can start creating and curating terms. As described before, the term editing forms are particularly easy to use and include all the property constraints that have been defined in the schema.

Fig. 3 shows the term editing form for a MADS/RDF `PersonalName` vocabulary whereby the property `madsrdf:identifiesRWO` takes as value a nested resource (of Class `madsrdf:Identifier`) that in turn includes several properties (fields) and Fig. 4 shows the

term page that includes the term URI, as well as all the terms' properties and values. URI references appear as hyperlinks that link to other resources. The user can also see the RDF representation of the term, either as RDF/XML or as JSON-LD serialization.

4 Access, re-use and integration mechanisms

Below we present the different channels through which vocabularies or individual terms hosted in Semantics.gr can be searched online and served for use or integration in third party systems.

A. Linked Data endpoints: A published vocabulary is publicly accessible as Linked Data. The RDF/XML and J-SON LD representations are available for every vocabulary or term, in the vocabulary's page or resource's page, respectively. Every vocabulary and term has a unique and permanent HTTP URI that returns its description in three different representations via content negotiation: a. in html (from a common browser or if the header parameter "Accept" has the "text/html" value), b. in RDF/XML (when the header parameter "Accept" has the "application/rdf+xml" or "application/xml" value) and c. in JSON-LD (when the header parameter "Accept" has the "application/ld+json" or "application/json" value).

B. Integration of predefined vocabularies within third party systems' cataloging forms via a Web API: Advanced re-use options are offered to authorized institutions that can use the available Web API to recall vocabularies from their digital content management systems (cataloguing systems, repositories, e-publishing platforms, etc.). Semantics.gr can be configured to populate targeted metadata fields with a set of controlled values from a particular source vocabulary on-the-fly. The APIs can be used in a cataloguing/archiving form to enable the user to insert values from controlled terms that pertain to a particular vocabulary published in Semantics.gr.

C. Mediated Resource Assignment Service: creation of a new resource and URI assignment service upon request: This mechanism allows a 3rd party registered system to request via a RESTful API and receive *asynchronously* from the curators of a vocabulary the URI of a vocabulary term⁸ that corresponds to an individual entity. The request is triggered automatically from within the cataloguing form he/she uses. The request's body embeds provenance information (registered repository, record, metadata field) and a description of the entity (person, concept, location etc.) to be assigned with a term from one or more predefined vocabularies. The description of the entity can be in a key-value format or, preferably, as RDF/XML.

The requests are allocated to authorized vocabulary curators through an in-built ticketing system. A curator receives the request in the personalized Mediated Resource Assignment Service dashboard. To process it, first he/she checks if there are terms matching the entity description. The tool automatically suggests matches based on

⁸ We use the word *term* for any entity/resource part of a vocabulary or catalogue, even for persons/legal bodies

string similarity on labels and on identifiers. If the curator confirms a match, he/she can easily enrich the matching term with additional information from the entity description using a merging assistant form. If there is no match, the curator creates a new term on the spot. Note that if the request embeds the description of the entity in RDF/XML and in the exact same data model as the vocabulary's, then the form for creating a new (or merging/enriching an existing) term is automatically filled with values obtained by the RDF/XML representation. The description of a person entity, for example, may include the preferred and alternative names, ORCID id and other identifiers, affiliation etc. This significantly eliminates the need for new data entry making the entire process faster. In either case, upon resolving a request the 3rd party registered system is notified via a webhook in order to handle the assignment appropriately (for example to enrich the metadata record with the term URI).

This service provides a simple and fast mechanism for 3rd parties that wish to integrate LD in their systems but need an additional curation step from an expert who checks for them if a term already exists in the designated vocabulary or creates a new term when needed. This is also beneficial for the institution that owns the vocabulary since it facilitates the expansion of a vocabulary in a controlled and curated way.

D. A Mapping Tool for enriching content *en masse*: This tool can be used to define mappings 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 mappings from source datasets (resources/terms from vocabularies, metadata records or aggregated metadata values) to terms from a target vocabulary. The tool incorporates a self-improving automatic suggestion mechanism and additionally supports the curator when intervening in the mapping process. After mappings for a dataset are finalized they can be served on request via a RESTful API in JSON format which can be used by the aggregator or repository to enrich the actual dataset easily and *en masse*. The tool is thoroughly described in [1].

5 How EKT leverages Semantics.gr to enrich and integrate its content

Until today there are 48 Vocabularies published on Semantics.gr covering different various thematic areas and scientific topics. Most of these vocabularies have been developed by EKT's staff and by EU-funded projects' consortia. The majority are original Vocabularies while some are adaptations and extensions of existing popular vocabularies with labels translated in Greek.

Vocabularies currently hosted in Semantics.gr are grouped around four categories based on their area of application: i) Persons/Corporate bodies (supported models: MADS/RDF PersonalName/CorporateName and edm:Agent) ii) Historical periods (supported models: edm:Timespan from EDM), iii) concepts (supported models: SKOS) and iv) Geographical locations (supported models: edm:Place from EDM).

We use the vocabularies that we have developed in-house to enrich our scientific and cultural content repositories and aggregators with Linked Data resources. These

activities have already significantly improved the quality and the efficiency of our content cataloguing, have boosted semantic interoperability and data integration across our distributed infrastructures and have improved the overall user search and browsing experience in our content discovering websites.

So far, we are using Semantics.gr in the following content infrastructures that are maintained by EKT:

- **SearchCulture.gr** | <https://www.searchculture.gr>: the national cultural data aggregator (more than 720K+ items/ records from 72 Greek institutions)
- **OpenArchives.gr** | <https://www.openarchives.gr>: the biggest Greek aggregator of scientific literature (770K+ items records, such as scholarly publications and gray literature, from 78 institutions)
- **National Archive of PhD Theses** | <https://www.didaktorika.gr>: collects in digital form doctoral dissertations awarded by Higher Education Institutions (HEIs) or to Greek scholars by foreign HEIs (46K+ thesis).
- **ePublishing** | <https://epublishing.ekt.gr>: EKT's open access scholarly communication service for the Greek academic community (20K+ articles).
- **Ariadne** | <http://ariadne.ekt.gr>: EKT Institutional Repository (3K+ items)

Next, we highlight our most representative vocabularies that are used in the above infrastructures:

Cultural heritage item types vocabulary⁹: A SKOS-based original vocabulary using the skos:Concept class to describe different types of cultural artifacts. It is hierarchical, bilingual and the majority of the terms are linked to the Getty Art and Architecture Thesaurus. The vocabulary is being used to enrich SearchCulture.gr collections as per item types using the *Mapping Tool* [1] (see Section 5, D).

UNESCO Thesaurus (EKT version)¹⁰: Vocabulary adapted from the UNESCO thesaurus. We followed the original hierarchical thesaurus structure whose concepts are grouped in 7 broad thematic areas. For the EKT version, 1387 terms were selected that are particularly suitable for the SearchCulture.gr collections. The vocabulary was used to enrich SearchCulture.gr items as per their subjects using the *Mapping Tool* [1].

Greek historical periods¹¹: A vocabulary constructed according to the semantic class edm:Timespan of Europeana's EDM model. It contains 94 terms that cover Greek history from 8.000 BC to today. It is hierarchical, bilingual, it covers the Greek territory and some values in the 3rd and 4th level correspond to individual civilizations. The vocabulary is used to enrich the content in SearchCulture.gr as per historical periods using the *Mapping Tool* [1].

Significant Figures in Greek History and Culture¹²: A vocabulary of approximately 7,500 names of historical figures and influential persons from across Greek history and culture conforming to the edm:Agent class of Europeana's EDM. Each entity is assigned a unique URI, enriched with basic biographical information and links to other reputable

⁹ <https://www.semantics.gr/authorities/vocabularies/ekt-item-types>

¹⁰ <https://www.semantics.gr/authorities/vocabularies/ekt-unesco>

¹¹ <http://semantics.gr/authorities/vocabularies/historical-periods>

¹² <http://semantics.gr/authorities/vocabularies/searchculture-persons>

databases such as VIAF. It is also enriched with links to a hierarchical bilingual vocabulary of professions/occupations (e.g. politicians, clergy, artists, scientists, etc.) that we developed for this purpose. The vocabulary, which is not published yet, is being currently used to enrich SearchCulture.gr items with creators or persons they are related to. The *Mapping Tool* is also used for this type of enrichment.

EKT Scientific Disciplines¹³: we have also developed a vocabulary of scientific disciplines based on the OECD 2015¹⁴. It aligns with the FORD classification with regards to the six 1st-level broad thematic areas and forty-two 2nd-level thematic areas. EKT information scientists expanded the 2nd-level thematic areas creating a 3rd fine-grained level based on a series of valid scientific resources. The resulting SKOS vocabulary comprises of 474 unique bilingual subject terms covering the principal areas of Science, Technology & Development. The vocabulary is being used in the National PhD Archive repository where PhD holders select the subject areas that their dissertation relates to. The vocabulary will be further used for classifying the content of OpenArchives.gr and ePublishing.

Catalogue of Authors of Scientific Publications: A list of authors from across EKT's scientific repositories such as the National PhD Archive repository, ePublishing and Ariadne. The vocabulary conforms to the MADS/RDF ontology (madsrdf:PersonalName class) and currently consists of more than 53k authors, initially imported from the National PhD Archive.

The screenshot displays the 'Mapping Tool' interface. At the top left, there is a 'Submit' button and a summary of mapping statistics: 'Total number of items: 67', 'Pending: 40', 'Will not be mapped: 0', 'Validated: 26', 'With suggestions: 1', and 'Under validation: 1'. The main area shows a mapping between two entities. The source entity is 'ejournal-periphery/28686_31458_31857_36641_46257_50166' with the name 'Χρυσομαλλίδης, Χαράλαμπος'. The target entity is 'persons/1761239616' with the name 'Χρυσομαλλίδης, Χαράλαμπος'. Both entities have a 'Mapping status' of 'Validated'. The interface also shows a list of 'Possible email' addresses and 'Field of activity' for both entities. A blue button at the bottom right says 'Update vocabulary resources according to this mapping'.

Fig. 5. Mapping authors from an ePublishing journal to the Catalogue of Authors of Scientific Publications

¹³ <https://www.semantics.gr/authorities/vocabularies/EKT-voc>

¹⁴ OECD FORD Research and Development classification fields. <https://unstats.un.org/unsd/classifications/Family/Detail/1039>

Our aim is to enrich the publications' records with the authors and contributors' URIs and to integrate a person's scientific and research output from across a range of different platforms. For new records we will integrate our scientific repositories with the *Resource Assignment Service* (described in Section 4, E). The National PhD Archive is the first platform fully integrated with this service allowing PhD theses records to get enriched with Person URIs from the catalogue prior to publication. For existing content, we are using the Mapping Tool to map author names to persons from the Catalogue and to enrich the Catalogue with new persons when there is no match. We have started this process with the authors of articles of selected journals from the ePublishing platform. Fig. 5 illustrates the mapping form of the tool, where authors from a specific journal (left part) are mapped to authors from the Catalogue (right part).

6 Related Work and Discussion

There are many vocabulary publishing tools, both open source and commercial. Most of them support only SKOS [2][5] or a limited pre-defined set of data models [3]. There are also platforms and research tools that support ontology and Linked Data resource management, such as VocBench [4] and Protégé [6] that focus more on ontology engineering leveraging OWL expressiveness for building ontology-based systems and are rather geared towards advanced users. These systems are also enabled with reasoning and SPARQL endpoint services confirming the consistency and information retrieval functionality of data models.

Semantics.gr on the other hand is not meant as an ontology engineering tool nor has it been purposed as a Triplestore service. It puts at the centre the development of a vocabulary network to be used by interested parties and therefore the building of a community network. This is done through an easy vocabulary creation and curation workflow that respects the reusability principles of the Semantic Web and a user-friendly GUI configuration mechanism that enables users to register and configure standard and custom data models and/or application profiles. The innovative trait about it is that it provides a central and uniform environment for developing and housing a vocabulary bank to be freely used by anyone who is interested in applying vocabularies to their collections. It can create dedicated forms for updating or creating vocabulary terms that conform to a schema, which in turn reflects or specializes the specifications of an established or custom data model. Moreover, Semantics.gr embeds advanced mechanisms that enable external platforms such as repositories and aggregators to link and align their content with respect to vocabularies, taxonomies, thesauri and authority files. This can be achieved either for newly created content via a Web API or a mediated resource assignment service or for existing content via an advanced mapping tool that embeds semi-automatic workflows (see Sections 4, B, C and D). All these not only set Semantics.gr as a core vocabulary publication service but also align with its wider objective to develop an active community around it.

7 Conclusions

With the development of Semantics.gr, it is the first time at national level in Greece, that state-of-the-art semantic knowledge representation technologies and methodologies have been applied to such a scale in scientific and cultural data infrastructures of national importance, for the homogenization, standardization, enrichment and integration of information resources.

The key innovation of Semantics.gr lies in that it enables even non-experts in semantic web technologies to harness the benefits of the Semantic Web to publish authoritative Vocabularies as Linked Data. The publication and re-use of Linked Data Vocabularies in digital infrastructures enhance the quality and interlinking of resources as well as the interoperability of the systems, while improving search and browsing functionalities, adding multilinguality and helping lower the costs of documentation. EKT aspires that Semantics.gr develops into a central public platform, where authoritative semantic vocabularies can be curated and published for everyone to re-use in Greece and beyond, building an active community around vocabularies. In the new, open linked environments these semantic tools can provide a powerful conduit for research and discovery of digital resources.

Acknowledgments

The work presented in this article has been partly 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.

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 multilinguality in the Greek cultural aggregator SearchCulture.gr. (2018) *Int. J. Metadata, Semantics and Ontologies*, Vol. 13, No. 1, pp.75–92
2. Suominen, O., Ylikotila, H., Pessala, S., Lappalainen, M., Frosterus, M., Tuominen, J., Baker, T., Caracciolo, C., Retterath, A. (2015). Publishing SKOS vocabularies with Skosmos. Manuscript submitted for review, June 2015
3. TemaTres: <https://www.vocabularyserver.com/index.html>
4. Stellato, A., Fiorelli, M., Turbati, A. et. al: A collaborative Semantic Web editor for ontologies, thesauri and lexicons, *Semantic Web*, doi:10.3233/SW-200370, 1-27, 05, 2020
5. Tuominen, J., Frosterus, M., Viljanen, K., Hyvönen, E. (2009). ONKI SKOS Server for Publishing and Utilizing SKOS Vocabularies and Ontologies as Services. In *Proc. of the 6th European Semantic Web Conference (ESWC 2009)*, vol. 5554.
6. Musen, M.A. The Protégé project: A look back and a look forward. *AI Matters*. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. DOI: 10.1145/2557001.25757003.