

Approaching Cross-Domain Search in Environmental Applications – Towards Linked Data

Gerald Schimak, Luca Petronzio, Tomas Pariente Lobo

▶ To cite this version:

Gerald Schimak, Luca Petronzio, Tomas Pariente Lobo. Approaching Cross-Domain Search in Environmental Applications – Towards Linked Data. 10th International Symposium on Environmental Software Systems (ISESS), Oct 2013, Neusiedl am See, Austria. pp.71-84, 10.1007/978-3-642-41151-9_8. hal-01457504

HAL Id: hal-01457504 https://inria.hal.science/hal-01457504v1

Submitted on 6 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Approaching cross-domain search in environmental applications- Towards Linked Data

Gerald Schimak¹, Luca Petronzio², and Tomas Pariente Lobo³

¹AIT Austrian Institute of Technology GmbH, Vienna, Austria gerald.schimak@ait.ac.at ²Telespazio Spa, Rome, Italy luca.petronzio@telespazio.com ³Atos Spain, Madrid tomas.parientelobo@atosresearch.eu

Abstract. TaToo – Tagging tool based on a semantic discovery framework, a project funded by the European Commission provides a web-based solution for easy and accurate discovery as well as tagging of environmental resources. The novelty relies on a semantic framework integrating different domain ontologies in a multi-domain and multilingual context. The underlying ontology framework, comprises besides the different domain ontologies (e.g. related to climate change, agro-environmental and anthropogenic impact domains) also concepts and methods to establish a mapping between the domain ontologies and socalled minimal environmental resource model (MERM). Together this forms a suitable and usable bridge ontology allowing a cross-domain discovery by using aligned ontologies concepts from different domains. The clear advantage for the end-user is that he is now able to find relevant information stemming from other domains, (like from impact of pollutant, climate change or temperature on human health) that he would not have found before but would be even more important to him that the ones only from his domain of expertise. The TaToo semantic framework extends cross-domain search evolving towards the Linked Data initiative by providing a linking functionality. The cross-domain search can be extended including in the search results also linked resources.

Keywords: semantics, information enrichment, cross-domain, search, discovery, annotation

1 Introduction

Today's Holy Grail is the development of cross boundaries (cross-domain, crossorganisations) usage of information, especially including search & discovery, interpretation, access, tagging, reliability, trustworthiness plus all the necessary steps needed of information enrichment for a broader usage of information objects.

The appeal of cross-domain searching is obvious. A simple example: Why, after all, should the text of, say, Shakespeare's Hamlet only be retrieved alongside other texts – an artificial constraint? Why not alongside related resources from other do-

adfa, p. 1, 2011. © Springer-Verlag Berlin Heidelberg 2011 mains such as sound recordings of performances of the play or digital images of the costumes?

Why not allowing someone else, e.g. a third party, to give his impression on an information object or to link it with some other information items of importance to him or to his/her community. This would really pave the way to a broader and better understanding of information often given by highly specified experts and improve its usage.

The problem so far can be described as a tragedy – as discussed in Havlik's "Tragedy of meta-information" [1] and also identified by ORCHESTRA (http://www.euorchestra.org/), an FP6-Integrated Project. It is that the "appropriate" metainformation depends on the intended use, which in turn can never be known in advance. For example, the set of meta-information required by a scientist interested in climate change may be very different from the meta-information required for automatic report generation in the context of "Clean Air For Europe" (CAFE) directive. Thus hundreds or thousands of meta-data structures have been defined and put in place.

Macolm Polfreman's article [2] brings it to the point: "The problem is that the archive, library, gallery, and museum communities have developed frequently incompatible metadata structures of their own". But it does not make sense just to blame on them as nearly everyone on the globe did it that way (and more or less it was discussed as good practice to provide meta-data).

Some promising approaches like the Dublin Core are widely used but have some weakness in loosing potentially information by their tight structuring of meta-information [3] and its involved mapping process.

So we had the idea to overcome all this tragedy of meta-data definition, acquisition and burden of collecting and entering them, by developing a semantic based system that allows to identify e.g. similarity by using ontologies going hand in hand with producing a way of cross-domain processing (searching, discovering, tagging) of information pieces (e.g. understood as resources in our naming conventions, which are data, services, models, etc.) and let doing the work of information enrichment (more or less the former entering of meta-information) peoples, experts, or communities interested in a specific topic touching various domains (e.g. like climate change or air pollution or its impact on human health).

In the TaToo project (Tagging Tools based on a Semantic Discovery Framework, Contract Nr. 247893; www.tatoo-fp7.eu) we stressed this issue specifically, and developed a semantic framework that allows the alignment of ontologies of different domains (like human health, agro-environmental, climate change) up to a bridge ontology (covering the necessary and minimal concepts suitable for the environmental domain). Following this approach we have been able to facilitate cross-domain search on the one hand and on the other to allow integration of additional domains respectively their domain ontology. This last point is still a piece of hard work, but at the end we can provide an added value by extending the cross domain search to the user, as a key service, that can enrich users experience significantly.

2 The Semantic Challenge

As introduced before TaToo's goal is semantic tagging and searching of environmental resources. TaToo aims to capitalise on the principles of the Semantic Web using ontologies as the underlying model for tagging and searching resources. The main objectives of the ontologies within the TaToo Framework are the following [4]:

- Allow formal tagging and searching of environmental resources;
- Allow contextual cross-domain tagging and searching of environmental resources;
- Foster multilingualism issues;
- Implement an extensible and more accurate searching mechanism.

To achieve this objectives one of the major challenges was the development of these ontologies. TaToo decided to base the development applying existing standards (W3C standards) and (re-)use proven methodologies like the NeOn methodology [5] and tools such as the NeOn Toolkit [6] as much as possible. But also to avoid the use of complex semantic for tagging and discovery.

2.1 The TaToo Ontology Framework

One can easily imagine a key challenge when applying semantics is the conceptualization and formalization of ontologies. As mentioned above the NeOn methodology for engineering ontologies was adopted [7]. This methodology provides guidance for all aspects of the ontology development, focusing on collaborative aspects on designing ontologies, the reuse of existing resources (ontological and non-ontological), and the further maintenance and evolution of the ontologies. In order to find out the common parts (terms, expressions, vocabulary) across the different domains (climate change, health, agro-environmental, etc.) covered by TaToo in its Validation Scenarios [8-10] several workshops, for introducing and collecting competency questions, were held.

Thereof, the main conclusions drove towards the definition and implementation of a common minimal model to describe environmental resources and annotations, and a semantic infrastructure where different domain ontologies shall be relatively easy plugged in; furthermore three main issues (stemming from an extensive requirements analysis [11]) impacted on the shape of the TaToo Ontology Framework, namely: i) take into account multilingual aspects; ii) allow formal tagging and searching environmental resources; and iii) allow cross-domain search.

With respect to the multilingual approach of the ontologies The final decision made in TaToo was to use the RDFSchema metadata rdfs:label and to add multilingual metadata storage in the TaToo Ontologies (including domain ontologies) [12].

Formal tagging and searching of heterogeneous and disperse environmental resources implies to define an ontology resource model. For practical reasons formal tagging in opposition to informal tagging should use a common and well-defined minimal model. For this purpose the Minimum Environmental Resource Model (MERM) was introduced. MERM is the cornerstone of the TaToo Ontologies, meaning that all resources annotated within TaToo contain tags according to the MERM model.

This approach implies that common elements of the domain ontologies have been abstracted and mapped to a common and really simple ontology in order to allow semantic interoperability. Within TaToo this common ontology is the so-called Bridge Ontology.

A high level description of the TaToo Ontology Framework can be seen in Fig. 1 below:

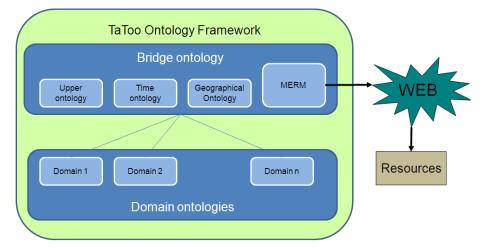


Fig. 1. High-level overview of the TaToo Ontology Framework

The ontologies defined in the Fig. 1 are as following:

- Minimal Environmental Resource Model (MERM): MERM is the ontology that gives structure to the annotations of a given resource, hence an effort to identify a minimal model that, without limiting the expression of specific domains, acts as a reference for past and future applications in the domain.
- MERM specifies concepts and properties that allow the description of what is important to say about a resource common to the environmental domain. MERM is extensible in order to allow specialization on different sub-domains, if needed.
- TaToo Bridge Ontology (BRIDGE Ontology): The BRIDGE Ontology main objective is to map elements coming from different domains fostering easier cross-domain integration. It acts as a bridge or hub where elements of different domains are mapped together. The BRIDGE Ontology also provides a unified way to annotate entities common to most of the sub-domains of the environmental domain. In this sense, the BRIDGE Ontology contains MERM plus a set of common domain elements, such as time and geographical ontologies. Domain ontologies can map to these common elements in order to achieve cross-domain semantic interoperability.
- Domain ontologies: A set of ontologies from the different sub-domains that are mapped to the common elements of the BRIDGE Ontology.

For the conceptualization of MERM three main classes are defined: Resource, AccesInfo, and Annotation.

- Resource is the main class for MERM. The Resource class includes annotation management information (author, date of creation, etc.). A resource presents a set of annotations describing what the resource is about and annotations used to discover the resource;
- ResourceAccessInfo is about accessing the resource. We assume that a resource has a single AccessInfo. The content of the access information depends on the type of the resource. Although access information for a web document would be different from access information for a web service, in TaToo Ontology we assume a single way to access to any resource; through its URI;
- Annotation describes what a resource is about according to the people tagging the resource. It is therefore an opinion about the resource, which might or not be true. The content of an annotation depends on the type of the resource (an annotation for a web document would be different from an annotation for a web service). Saying that "some resource is related to some domain concepts" would be the most basic type of annotation in TaToo. Resources can be tagged and discovered taking annotations as basis;
- Evaluations are also part of the MERM model, allowing users to evaluate both resources and resource annotations according to various evaluation criteria;
- Linking part of the MERM model enables TaToo resources to be linked by two types of links. The first type refers to the so-called typed links defined by a selected set of well-established properties such as rdfs:seeAlso, rdfs:isDefinedBy, dcterms:hasPart, dcterms:isPartOf, etc. The second type refers to similarity links that link two similar TaToo resources. The MERM model introduces its own property (merm:similarTo) that describes the similarity links.

As an implementation convention, within TaToo all domain elements subject of annotation must be subclasses of the bridge:Topic class.

The BRIDGE Ontology imports MERM, plus the environmental BRIDGE elements (GeoNames [14], and NUTS ontologies W3C Time [15]).

More details on the relation between MERM annotations and domain elements is managed using some BRIDGE Ontology elements can be found in [4].

2.2 Procedure to add new ontologies to the TaToo Framework

There is no single standard ontology or shared vocabulary that encompasses the entire environmental domain, so the TaToo Ontology Framework faced the challenge of being capable of dealing with annotations from different domain ontologies while enabling a certain degree of cross-domain discovery capabilities, and at the same time offering a unified, extensible and relatively simple ontology framework.

The main purpose of the TaToo Domain Ontologies is that of offering new domain elements (individuals) as topics (individuals of subclasses of bridge:Topic) to allow semantic annotation and search of resources following the MERM model. This implies that the focus of TaToo is not only in interoperability issues, but also of allowing

a simple an effective way of adding new domain elements for the annotation and discovery, taking into account cross-domain considerations.

Cross-domain mappings are the most complex issue when adding a new ontology. Allowing direct mappings between different domain ontologies could potentially endup in a really unmanageable network of interconnected ontologies. To avoid this issue, we propose to abstract the mappings in a way that all classes or individuals subject of mapping will be available at the bridge ontology. The requisites for introducing a new domain ontology into TaToo are the following:

- The new domain ontology must be an OWL2 RL ontology [13];
- The classes of individuals that can potentially be used for annotation (topics) must be made explicitly subclasses of the bridge:Topic class of the bridge ontology;
- The domain ontologies must provide individuals of the topics (subclasses of bridge:Topic) that will be used for annotation. TaToo does not allow annotation of resources using classes in order ease the inference process and get an optimal performance.

3 Information Enrichment implementation

The TaToo Framework enables experts as well as arbitrary users to share trusted and reliable information and to allow easy discovery and semantically enhanced tagging [16] of existing environmental resources.

The TaToo Framework has gone through several short-end software lifecycles to progressively design and implement the core business logic, the extension of functionalities and finally the consolidation of the framework together with the enrichment of available tools and integration of Linked Data initiative. The framework demonstrates also the use of the ontology framework (as already discussed above) engineered during the TaToo project as a parallel implementation activity and how its integration within the implemented software framework enables cross-domain search and tagging, Linked Data extension, similarity between resources and multilingual aspects. Having in mind also to design a flexible and robust system the TaToo Framework architecture is composed of four different tiers:

- Presentation tier: provides tools to the end users to access the TaToo Knowledge Base;
- Public Service tier: provides Web Services interfaces to directly access the framework functionalities;
- Business tier: implements the core business logic of the framework, in particular semantic reasoning and harvesting;
- Cross tier: includes security and administration services.

3.1 Presentation tier

The Presentation tier contains all the components the user can take advantage of to access the functionality provided by the TaToo Framework, in particular the tagging, discovery, evaluation and geo-referentiation functionalities.

The implementation of a portal [17] as a first entry point and showcase to access the TaToo Framework has been motivated by the TaToo vision of providing easy access to users, thus only requiring a browser and an Internet connection wherever the users are located, with no requirement to the user to install other software or plug-ins.

The reference implementation of the TaToo Portal is based on Liferay Community Edition technology (http://www.liferay.com/).



Fig. 2. TaToo Web Portal

The implementation of the final version of the TaToo Portal have been influenced by several live and online evaluation and validation surveys performed by end users that have been conducted during Scientific Conferences or by invitation sent by the TaToo Project Consortium.

3.2 Public Service tier

The Presentation tier provides the Web Services interfaces (SOAP) in order to perform semantic Tagging, Discovery, Evaluation and Linking operations on TaToo resources.

The Tagging Services expose the tagging functionality of the TaToo System through its public interfaces. The Tagging Services allow to:

- Register new resources;
- Associate annotations with resources;
- Retrieve annotations associated with resources;
- Delete annotations.

The Discovery Services expose the public discovery functionality to User Components. They support semantic search and discovery of annotated resources. The discovery process is query driven, allowing the user to selected certain terms from an ontology that are then used for the semantic search.

In particular the following search operations are available:

- search: This is the main discovery operation, allowing the client to filter by topic, by annotation types, and by several resource and annotation metadata according to the TaToo Ontology Framework. The client may also decide whether the search will also retrieve similar resources and/or decide if they would like to perform cross-domain search;
- rectangleSearch: Main operation for performing GeoNames location-based queries within a rectangle delimited by two GeoNames coordinates;
- nutsSearch: As in the previous case, this operation is intended to perform locationbased queries, but based on NUTS regions (from the NUTS ontology).

The Evaluation Services provide necessary functionalities for the evaluation of the TaToo-managed resources and the resource annotations. The functionalities of the Evaluation Services could be grouped into the following four categories:

- Creating new evaluations of the resources and the resource annotations;
- Retrieving existing resource and annotation evaluations;
- Removing existing resource and annotation evaluations, and;
- Generating and retrieving average evaluation scores of the resources and resource annotations.

The TaToo linked data functionalities are implemented through the Linking Services. In particular, they provide the following list of functionalities:

- Adding a predefined set of typed links between TaToo resources;
- Adding similarity links between similar TaToo resources;
- Retrieving a list of resources linked to a given resources;
- Retrieving a list of similar resources of a given resource;
- Removing the typed links and similarity links set between TaToo resources.

The Linking chain allows to "link" resources by a predefined set of properties available from the TaToo Ontology framework.

3.3 Business tier

The Business tier is composed of the following set of components [18]:

- The Semantic Processor;
- The Harvester and Connectors, and;
- The Tagging, Discovery, Evaluation and Linking processors (where the business logic of operations resides).

The Semantic Processor is a Building Block made of a set of Core Components. It offers the infrastructure to access to the TaToo Semantic Repository and ontologies allowing semantically enhanced tagging, discovery, evaluation and linking functionalities.

The Semantic Processor provides functions to:

- Retrieve ontologies on the basis of user-defined properties like domain or context;
- Store and retrieve semantic annotations for resources (RDF-Triples);
- Search for semantic annotations;
- Semantic reasoning based on the existing ontologies.

The semantic functionality is realised by the following components composing the Semantic Processor internal architecture:

- The Ontology Manager to manage access to the TaToo Ontologies;
- The Semantic Framework to manage the interaction with the Semantic Repository and the Reasoner;
- Reasoner to infer new knowledge from the available RDF-Triples & ontologies and to check for incongruent (or inconsistent) information while managing ontologies;
- Triplestore (Semantic Repository) which is a data base for the storage and retrieval of RDF-Metadata.

The Resource Harvester is the Core Component responsible for harvesting metainformation from available resources that could be data (catalogues), web services, web pages, etc. A more and more large meta-information set is essential to improve the process of searching for resources returning valuable results to the requesting user.

As result of the final implementation phase the following Harvester connectors have been implemented and are currently harvesting resources via the:

- GENESI-DEC Connector harvests OGC meta-information from GENESI-DEC catalogues, only those were pertaining to TaToo domains are then stored in the repository;
- JRC Connector harvests meta-information from JRC Validation Scenario catalogues, generated in the context of the TaToo project;
- Masaryk University Connector harvests meta-information from Masaryk University Validation Scenario, generated from the legacy RDF;
- Linked Data Connector harvests resources pertaining to the TaToo domains that are part of Linked Data;
- Web Site Connector harvests meta-information stored as RDFa in any Web Site available. As a proof of concept the TaToo Project Web Site has been harvested and stored as a resource.

The Linked Data Connector provides harvesting capabilities to the TaToo Framework to crawl and harvest Linked Data resources. The connector is able to integrate TaToo resources annotated in the Knowledge Base repository with linked resources from the Semantic Web. The connector presents then two main advantages, largely increases

the knowledge base stored in the TaToo repository, and plugs the TaToo resources in the Linked Open Data cloud.

The Linked Data Connector is based on an already existing web crawler adapted to traverse and harvest content from the Linked Data web. The first action of the Linked Data connector harvesting workflow is then to crawl resources from the Linked Web. The implemented connector filters the result triples based on the link types supported in TaToo. Once the Linked Data connector receives the filtered set of triple statements, it directly stores the triples set using the TaToo Linking library.

The final step of the Linked Data workflow is to perform a second round parsing of the harvested LOD (Linked Open Data) cloud datasets to check if there are meaningful annotations in the context of TaToo. In this case a specific mapping is performed and resources are also semantically annotated with concepts from TaToo Domain Ontologies.

3.4 Cross-domain, similarity and multilinguality samples

The aim of this section is to provide a few views on how using the TaToo Portal as an entry point and how it is possible to obtain results from cross-domain searches, or extend the semantic search results by using Linked Data and similarity, and finally how to demonstrate how multiinquality got implemented.

The Cross-domain search feature is implemented in TaToo as optional, the framework leaves to the user the choice whether to perform a search only in its own domain or extend it to other domains.

The User can access the TaToo Portal and operate a search through the Search tools depicted in Fig. 3.

There the user has used a concept from the Regional Climate Change domain selected in the Ontology Tree above in the portal and then in the search results are displayed, properly ranked, resources from the Regional Climate Change domain together with resources from the Earth Science (OGC – Open Geospatial Consortium) domain, annotated with a different concept from the GENESI-DEC Domain Ontology that is aligned with the concept used to search from the Regional Climate Change Domain Ontology.

In the same search it was also selected the "Extension" option that triggers the usage of Linked Data. The search results include then together with the resources found thanks to their semantic annotation the linked resources, that means bound together by using one of the provided set of links, or properties, within TaToo.

Fig. 4 shows one of the linked resources found in the previous search.

	Η		57 Jal		Welcome	e TaToo Tools
191211111111						
aToo Portal	TaToo Tools					
Domain		erties Resource properties				
Domain	Type Annotation prope	erites Resource properties				
Select D						
	al Climate Change	•				
Select a	Topic			* Selected to Energy	pics	
			>	Litergy		
TOPICS		URI				
► Buil	iding nate Adaptation	http://www.ait.ac.at/ClimateTwins_Domain#Buile http://www.ait.ac.at/ClimateTwins_Domain#Clim				
► Ciin		http://www.ait.ac.avClimateTwins_Domain#Clim http://sweet.jpl.nasa.gov/2.2/guanEnergy.owl#Er				
	eorological Phenomena	http://sweet.jpl.nasa.gov/2.2/quanenergy.ow/#Er				
•		III				
						Search
						Search
	unter i contractione de la contractione de contractione de la contractione de la contrac					Search
esources	un en suu suu suu suu suu suu suu suu suu su		TTLE			Search
esources IRI		=metadata&series=NOAA:gov.noaa.class.GVAR_JI		ted Resource	e http://cwic.genesi-dec	Search
esources IRI ttp://cwic.gene	si-dec.eu/search?request		New Harves			
esources RI ttp://cwic.gene ttp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI	New Harves New Harves	ted Resource	http://cwic.genesi-dec	.eu/search?request=me
esources RI ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA:gov.noaa.class.AVHRR	New Harves New Harves New Harves	ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec	.eu/search?request=me .eu/search?request=me
esources IRI Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=meladata&series=NOAA:gov.noaa.class.GVAR_JI =meladata&series=NOAA:gov.noaa.class.AVHRR =meladata&series=NASA.FIFE_SF_BL92	New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec	:eu/search?request=me :eu/search?request=me :eu/search?request=me
esources RI ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=meladata&series=NOAA.gov.noaa.class.GVAR_JI =meladata&series=NOAA.gov.noaa.class.AVHRR =meladata&series=NASA.FIFE_SF_BL92 =meladata&series=NASA:s2k_mongu_irradiance	New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec	.eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me
esources RI ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA:gov.noaa.class.AVHRR =metadata&series=NASA:FIFE_SF_BL92 =metadata&series=NASA:s2k_mongu_irradiance =metadata&series=NASA:s2k_mondus_mod05_war	New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec	eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me
esources RI ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene ttp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA.gov.noaa.class.AVHRR =metadata&series=NASA:FIFE_SF_BL92 =metadata&series=NASA:s2k_mongu_irradiance =metadata&series=NASA:s2k_modis_mod05_wa =metadata&series=NASA:s2k_modis_h2o_flux	New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec	eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me
esources IRI Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene Ittp://cwic.gene	si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request si-dec.eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA.gov.noaa.class.AVHRR =metadata&series=NASA:FIFE_SF_BL92 =metadata&series=NASA:s2k_mongu_irradiance =metadata&series=NASA:s2k_modis_mod05_wa =metadata&series=NASA:s2k_modis_h2o_flux	New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource ted Resource	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec Annotation details object	:eu/search?request=me :eu/search?request=me :eu/search?request=me :eu/search?request=me :eu/search?request=me :eu/search?request=me
esources ire ittp://cwic.gene ittp://cwic.gene ittp://cwic.gene ittp://cwic.gene ittp://cwic.gene ittp://cwic.gene ittp://cwic.gene	si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA.gov.noaa.class.AVHRR =metadata&series=NASA:FIFE_SF_BL92 =metadata&series=NASA:s2k_mongu_irradiance =metadata&series=NASA:s2k_modis_mod05_wa =metadata&series=NASA:s2k_modis_h2o_flux	New Harves New Harves New Harves New Harves New Harves	ted Resource ted Resource ted Resource ted Resource ted Resource TYPE Web Annota	e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec e http://cwic.genesi-dec Annotation details OBJECT Oritoge.gatepro=sean	.eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me
http://cwic.gene http://cwic.gene http://cwic.gene http://cwic.gene	si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request DESCRIPTION This annotation is a resu	=meladata&series=NOAA.gov.noaa.class.GVAR_JI =metadata&series=NOAA.gov.noaa.class.AVHRR =metadata&series=NASA.S2K_mongu_irradiance =metadata&series=NASA.s2K_mondus_mod05_wa =metadata&series=NASA.s2k_modis_h2o_flux =m	New Harves New Harves New Harves New Harves New Harves Author	ted Resource ted Resource ted Resource ted Resource ted Resource	http://cwic.genesi-dec http://cwic.genesi-dec http://cwic.genesi-dec http://cwic.genesi-dec http://cwic.genesi-dec http://cwic.genesi-dec OBJECT OTIGECAGEPTO=Sear This annotation is a http://cwic.genesi-dec bridgetaloharvester	:eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m
esources Intel//cwic.gene http://cwic.gene http://cwic.gene http://cwic.gene http://cwic.gene http://cwic.gene intel//cwic.gene intel//cwic.gene intel//cwic.gene intel//cwic.gene	si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request si-dec eu/search?request DESCRIPTION This annotation is a resu	=metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NOAA:gov.noaa.class.GVAR_JI =metadata&series=NASA:gov.noaa.class.AVHRR =metadata&series=NASA:s2k_mongu_irradiance =metadata&series=NASA:s2k_modis_mod05_wa =metadata&series=NASA:s2k_modis_h2o_flux =metadata&series=NASA:s2k_modis_h2o_flux =metadata&series=NASA:s2k_modis_h2o_flux	New Harves New Harves New Harves New Harves New Harves New Harves Author tatooharve	ted Resource ted Resource ted Resource ted Resource ted Resource TYPE Web Annota	http://cwic.genesi-dece http://cwic.genesi-dece http://cwic.genesi-dece http://cwic.genesi-dece http://cwic.genesi-dece http://cwic.genesi-dece http://cwic.genesi-dece Decention details OBJECT ONIGE addepto=sear This annotation is a r http://cwic.genesi-de bridgetatooharvester Ontology_GENESI_D	:eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .eu/search?request=me .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m .cu/search?request=m

Fig. 3. Cross-domain search results

Resources:		Resources found:		
ID D		LINK TYPE	RESOURCE URI	
http://en.wikipedia.org/wiki/Passive_solar_building_design		similarTo	http://www.energieburgenland.at/oekoenergie/	
http://www.energieburgenland.at/oekoenergie/windkraft/un				
http://www.energieburgenland.at/oekoenergie/windkraft/win	+ similarTo +			
۰ m	Reload	*	m	•

Fig. 4. Linked Resources search results

Finally, thanks to the Ontology engineering combined with the Software implementation best practices the multilinguality has been integrated within the framework at different levels:

- User Interface available in different languages;
- Ontology terms available in different languages;
- Tagging and Discovery using different languages.

Fig. 5 shows the Tagging Portlet translated in Czech language. Both User Interface elements and Anthropogenic Impact Domain Ontology are available in the selected language and can be used to search and annotate.

IOVAT ZDROJE	59615147647				
Značkovaný Zdroj					
http://en.wikipedia.org/wiki/Ovary_cancer			•	Informace o Zdroji	
Vyberte Vaši Doménu					
Anthropogenic Impact *				🔲 📢 Pokročilé Značkován	
Vyberte Téma					
Filtrovat doménová témata					
TOPICS		URI	Typ Zdroje		
▼ CancerTopic		http://ontology.genasis.cz#CancerTo *	Vyberte Typ	*)	
▶ Medicine		http://ontology.genasis.cz#Medicine	Titulek Poznámky		
▼ Skupiny pacientů		http://ontology.genasis.cz#PatientGr			
▶ Gender		http://ontology.genasis.cz#Gender	Komentář k pozná	mce	
P Goldan		http://ontology.genasis.cz#RiskFactc			
▼Rizikový faktor		http://ontology.genasis.cz#Smoking +			
		http://ontology.genasis.ca#smoking +			

Fig. 5. Multilinguality

4 Conclusion

In this article we stressed the importance as well as the challenges of a semantic framework to reach to a satisfying extent cross-domain search and discovery. Still there remains the burden of doing alignments of ontologies of the different domains to the common part of all of them which is the bridge ontology. A task that is not easily done in couple of minutes.

But, what is gained by the TaToo approach is that we can extend the system by allowing to include new domains respectably domain ontologies and we do not loose information, in the contrary we enrich information to an object by additional tagging using a well-defined environment. We can link and reference to other (similar) resources. All this facilitates interoperability and understanding of cross domain issues.

Note: The majority of the software (e.g. public services) developed by the project are open source and can be requested by the TaToo Consortium Partners. The project documentation (deliverables) is available on the TaToo WebPage (www.tatoo-fp7.eu).

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement Number 247893. The authors would like to thank all the Partners of the TaToo Consortium for their valuable inputs and successful execution of the project.

References

- 1. Havlik D., Schimak G.: The Tragedy of Meta Information in: "Geophysical Research Abstracts", EGU General Assembly, Vol. 12 (2010)
- Polfreman M.: Cross-domain searching of library, archive and museum metadata some considerations (Museums and Images JISC-FAIR Cluster Group - Cross Domain Searching Issues Paper; Document Date: 1. July 2004)
- 3. Woodley M.: Crosswalks: the path to universal access. In: Introduction to metadata: pathways to digital information. Los Angeles, Calif.: Getty Information Institute, 2000. Available from: : http://www.getty.edu/research/institute/standards/intrometad
- ata/2_ articles/woodley/
 Pariente Lobo T., Fuentes J. M., Sanguino M. A., Yurtsever S., Dihe P., Nesic S., Avellino G., Petronzio L.: Semantic Framework Implementation Prototype, Deliverable 4.1.3 of TaToo project, Public Document, 2012 (Downloadable from www.tatoo-fp7.eu).
- Gómez-Pérez A., Motta E., Suárez-Figueroa M.C.: NeOn Methodology in a Nutshell. Ne-On-Book, http://www.neonproject.org/nw/NeOn_Book, (2010)
- 6. NeOn-Toolkit: The NeOn Toolkit, http://neon-toolkit.org/
- Suárez-Figueroa, M. C., et al. NeOn D5.4.1. NeOn Methodology for Building Contextualized Ontology Networks. February 2008.
- Kaufmann A., Peters-Anders J.: Refined Scenario Definition Case 1- V3, Deliverable 5.1.7 of TaToo project, Public Document, 2012 (Downloadable from www.tatoofp7.eu).
- Donatelli M., Fanchini D.: Refined Scenario Definition Case 2- V3, Deliverable 5.1.8 of TaToo project, Public Document, 2012 (Downloadable from www.tatoo-fp7.eu).
- Kubásek M., Hřebíček J.: Refined Scenario Definition Case 3- V3, Deliverable 5.1.9 of TaToo project, Public Document, 2012 (Downloadable from www.tatoo-fp7.eu).
- Schimak, G., Božić B., Kaufmann A., Peters-Anders J., Dihé P., Pariente Lobo T.: The TaToo Semantic Case – Requirements, impacts and applications. In: W. Pillmann, S. Schade, P. Smits (Eds.) Proceedings of the 25th EnviroInfo 2011:"Innovations in Sharing Environmental Observation and Information", pp. 832 – 847. SHAKER Verlag, 2 (2011.
- Dihe P., Avellino G., Petroncio L., Pariente T., Fuentes J. M., Yurtsever S., Rizzoli A., Nesic S., Božić B., Schimak G.: Semantic Service Environment and Framework Architecture V3, Deliverable 3.1.3 of TaToo project, Public Document, 2012 (Downloadable from www.tatoo-fp7.eu).
- 13. OWL 2 Web Ontology Language http://www.w3.org/TR/owl2profiles/#OWL_2_RL
- 14. GEONAMES: GeoNames ontology, http://www.GeoNames.org/ontology
- 15. OWL-Time OWL time ontology, http://www.w3.org/TR/owl-time/
- Avellino, G., Lobo T. P., Schimak G., Rizzoli A. E., Dihé P.: Applying semantics in the environmental domain: The TaToo project approach. In: W. Pillmann, S. Schade, P. Smits

(Eds.) Proceedings of the 25th EnviroInfo 2011:"Innovations in Sharing Environmental Observation and Information". pp. 848-855. SHAKER Verlag, 2(2011)

- 17. Avellino, G., Petronzio L., Lobo T. P., Lopez J. M. F., Yurtsever S.: The TaToo Portal Tagging and Discovery within TaToo. In: W. Pillmann, S. Schade, P. Smits (Eds.) Proceedings of the 25th EnviroInfo 2011:"Innovations in Sharing Environmental Observation and Information"; SHAKER Verlag, (2011)
- Dihé, P., Frysinger, S., Güttler, R., Schlobinski, S., Petronzio, L., Denzer, R., Nešić, S., Lobo, T.P., Schimak, G., Hřebíček, J., Donatelli, M.: An Architecture for the Semantic Enhancement of Environmental Resources. In Jiří Hřebíček, Gerald Schimak, Ralf Denzer (eds.). 9th IFIP WG 5.11 International Symposium on Environmental Software Systems: Frameworks of eEnvironment, ISESS 2011. pp. 372-384. Springer, Heidelberg (2011).