

The TaToo Portal – Tagging and Discovery within TaToo

Luca Petronzio¹, Giuseppe Avellino¹, Tomás Pariente Lobo², Jose Maria Fuentes², Sinan Yurtsever²

Abstract

A Web Portal represents an entry point on the Internet that allows users to access a centralized knowledge and information repository. In the TaToo project the Web Portal is the entry point to the Tagging and Discovery functionalities provided by the TaToo core framework. The aim is to supply a user front-end application which addresses user needs and that can be widely adopted across different domains and different locations. Moreover, a Web Portal benefits from the Web ubiquity increasing the number of possible users or the number of contacts from the same user. The users through the TaToo Web Portal are able to perform semantic searches for resources, to semantically annotate resources and to evaluate already existing resources or annotation associated with resources.

1. Introduction

The TaToo Framework architecture is composed by four different building block based tiers: the Presentation tier, the Service tier, the Business tier and the Data tier. In the Presentation tier the main entry point of the TaToo Framework, and also the reference implementation, is represented by the TaToo Portal. The motivation of implementing a Web portal as the access point to the TaToo functionalities is driven by the project philosophy to provide an easy access to the user, thus only requiring a Web browser and an Internet connection wherever the user is located. This means that to take advantage of Tagging and Discovery provided by TaToo it is not required to install locally any software or plug-ins.

The Web availability of the provided Graphical User Interface increases also the visibility of the TaToo Semantic framework, which is fundamental for a wide adoption of the TaToo tools. The TaToo Portal can then be defined as a central and user-friendly user-configurable entry point.

Portals are normally realised through the support of a portal server, and their functionality are provided by a set of modules named portlets. Portlets can be seen as independent TaToo tools delivering different functionalities, integrated within the TaToo Portal.

In the TaToo Portal three different types of portlets are identified depending on their nature: Tagging portlets, Search & Discovery portlets, Evaluate / Validate portlets. However, the concept of the TaToo Portal leaves the possibility to deploy and take advantage of any other type of portlet: administration, social, events, chat, and so on (see Figure 1).

¹ Telespazio S.p.A., Rome, Italy, email: {luca.petronzio, giuseppe.avellino}@telespazio.com

² Atos, Madrid, Spain, email: {tomas.parientelobo, jose.fuentesl, sinan.yurtsever}@atosresearch.eu

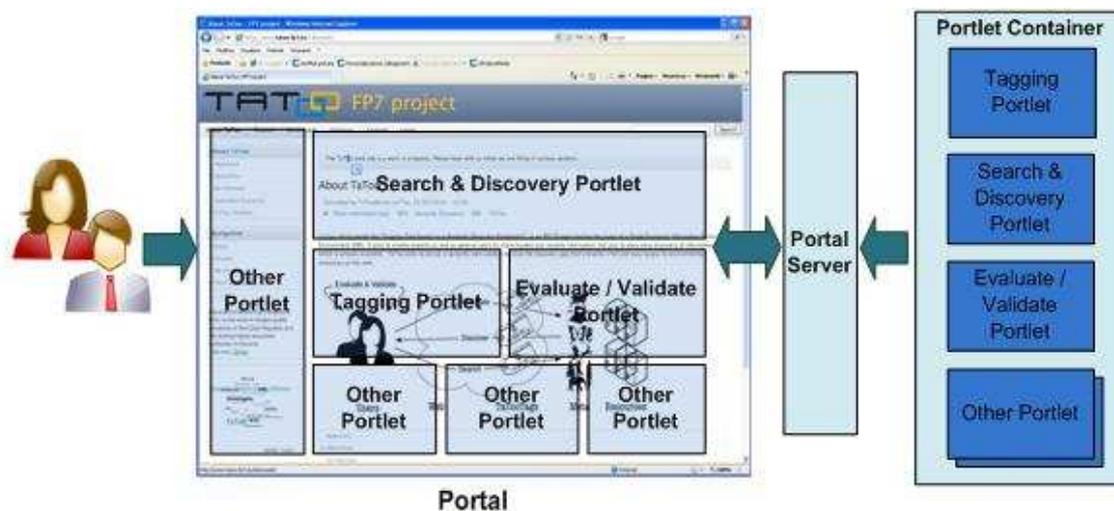


Figure 1 TaToo Portal

A further aspect of the TaToo portal which is strongly in favour of the end user is the personalization. In the TaToo Portal users can arrange their own pages, public and private, in the most conformable way that suits their needs. It is possible to deploy different portlets providing specific TaToo framework operations and setting up the working environment that the user was looking for from TaToo.

2. TaToo Portal

The TaToo Portal represents the main entry point for the TaToo semantic framework and its functionality. Due to the nature of a Web portal the user has access independently of his location to all the tagging, discovery, and evaluation / validation functionality.

In particular, the TaToo Portal provides:

- searching for semantically enriched resources;
- adding tags and semantic annotations to resources (discovered or directly provided by the user through an identifier – an URI);
- evaluating resources (possibly visualising them) and validating previously added tags and annotations.

The TaToo Portal, other than exposing TaToo core functionality as a reference implementation, can be customised to address particular requirements from different users from specific application domains. In particular TaToo is considering those domains from the TaToo Validation Scenarios, that are from the Masaryk University (MU), the Austrian Institute of Technology (AIT) and the Joint Research Centre (JRC).

The TaToo Portal is composed of a set of portlets that match with the set of services provided by the TaToo semantic framework. Portlets are developed independently of the portal itself, and are loosely coupled with the portal. Portlets are intended to deliver a single functionality. In case of customization or particular needs, more than one portlet can be offered concerning the same functionality.

The TaToo architecture identifies three basic portlets providing: tagging, search and discovery (results visualisation), evaluate / validate functionality (see Figure 2).

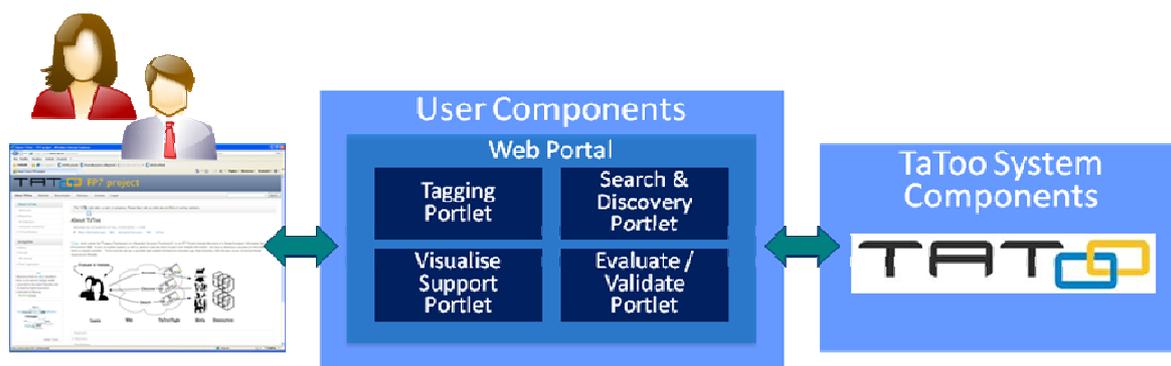


Figure 2 TaToo Portlets

In the TaToo project the Web Portal takes advantage of a particular open source portal server, Liferay. Liferay is the world's leading open source enterprise portal solution using latest Java and Web 2.0 technologies. Liferay portal choice has been driven by multiple considerations and characteristics, first of all it runs on all major application servers and web application containers, offering over than 700 deployment combinations. In TaToo two different setups have been tested: Glassfish and Tomcat application containers, supported by PostgreSQL version 9.0 as production environment database. After a careful evaluation Apache Tomcat 6 has been adopted in combination with Liferay 6.0.5 due to performance and scalability reasons. Liferay is compliant with the two portlet standards, JSP 168 and JSP 286. Liferay also provides a highly granular permission system that allows the administrator to customise user experience at organizational or personal level. Moreover, the Single Sign-On Authentication technology is supported with regard to the principal existing frameworks, in particular OpenID, Shibboleth, and CAS.

Portals also allow personalization, providing content aggregation functionality (content integration from different resources within a Web page). Personalization is one of the portal distinctive aspects that allows the user to create a comfortable working environment. The user can access its own portal control panel and arrange new private and / or public pages. In these pages the user can then select a set of portlets from different contexts, and add, move or delete them in order to take advantage of the required functionalities. Moreover, the TaToo Portal offers the possibility to deploy user developed portlets.

At present time the TaToo Portal is made of three pages referring to different thematic areas: 'Tagging', 'Discovery', and 'Masaryk University' (made of a customised User Interface for tagging resources). Of course the user can configure the portal as they like putting all the desired portlets in a single public or private user page or adding new user pages.

3. Discovery

As already stated, the TaToo portal provides a set of discovery portlets used for searching among the TaToo tagged resources. They provide user-friendly GUIs to access discovery functionality. This TaToo discovery functionality is accessible via the TaToo Discovery Services that make use of the underlying discovery core and the TaToo ontology framework. The core components implement the discovery business logic, while the ontology framework aims at providing a common knowledge foundation consisting on a set of different domain ontologies integrated in TaToo and some other common bridge elements. Within this knowledge representation layer, the TaToo bridge ontology contains ontologies such as the Minimal Environmental Resource Model (MERM) as the skeleton of the resource annotations, and other common ontologies widely used in the environmental domain (for time, geo-location, network of people, etc.). Be-

sides, the bridge ontology is used to generate cross-domain alignments between the different domain ontologies provided by the different scenarios.

This information space provided by the different domains is not only expected to evolve in time, but also new scenarios are likely to emerge and so new perspectives and domain knowledge representation needs. The integration process of these new models should avoid a great extent modification on already implemented portlets, minimising the effort on updating the generic portlets provided within the TaToo portal. Without a formalization of the interaction with the new models, the ideal of having ‘fixed’ services interfaces will fail. That is one of the main reasons why the MERM ontology is used as the structure to retrieve annotation information through the Discovery Service.

While interfaces are fixed, the extensible and modular nature of portals and portlets allows to easily integrate new perspectives that take advantage of the TaToo information models. The current implementation of the discovery portlets is domain agnostic, but scenario specific portlets within the TaToo Portal or other external clients can be developed in the future using the existing discovery service interfaces provided by the platform. So far, several discovery interfaces for expressing the search needs have been implemented: searching by name, choosing a category from a hierarchy or searching via facets with multiple parameters. There are also several ways of presenting results of the search to the user such as a table-like view or showing the retrieved annotations in a map or a faceted approach. Thus, several combinations of search and presentation portlets can be composed due to the configuration and portlet intercommunication capabilities provided by the portal.

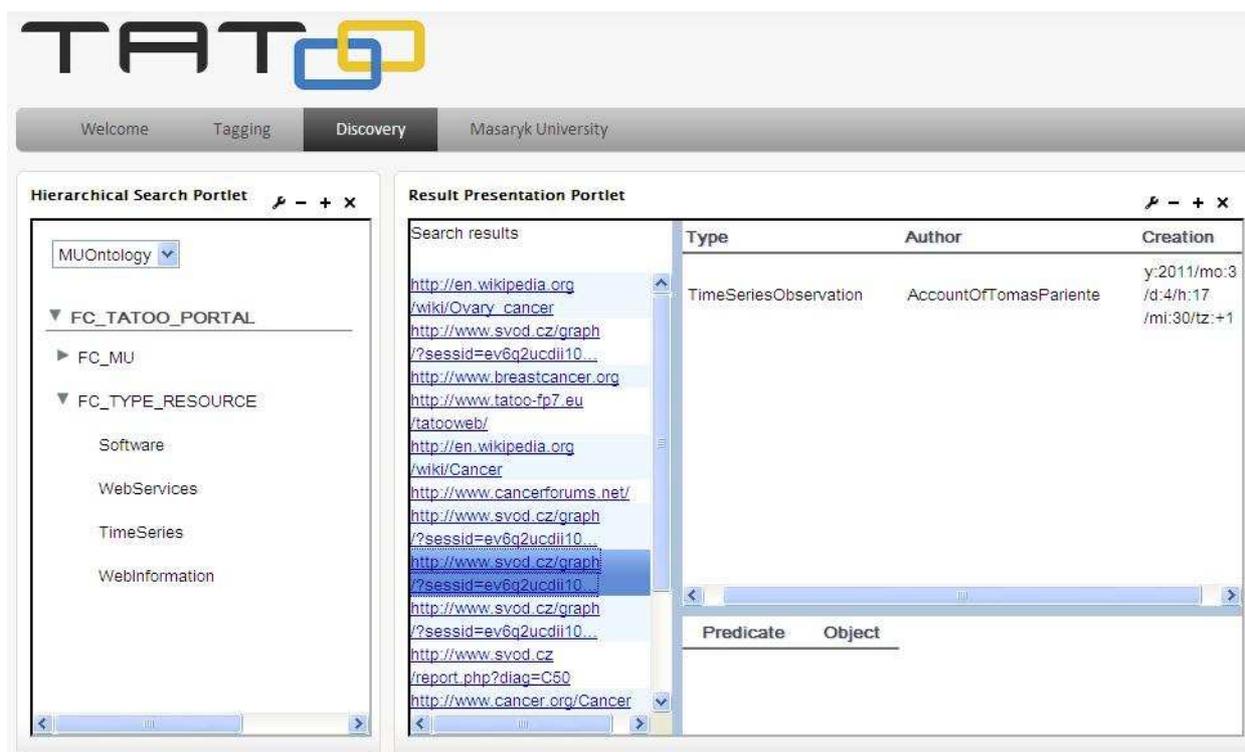


Figure 3 Hierarchical Search and Result Presentation portlets

Figure 3 presents a concrete example of these composition capabilities. Two portlets are deployed: the Hierarchical Search and the Result Presentation portlets. The Hierarchical Search portlet is the first implemented portlet to perform discovery of resources taking advantage of meta-information (tags) stored in

the TaToo knowledge base according to the existing ontology schemas. The portlet presents a tree-based view of the domain ontology where the user is able to navigate and select the desired concept to initiate the search process. The Result Presentation portlet shows the meta-information associated to the resources retrieved for the query. This Result Presentation Portlet shows a RDF-like (subject-object-predicate) representation of the tags associated to resource.

In order to improve the user-friendliness of the results presentation, the Faceted Result Portlet has been introduced. This portlet is based on Exhibit technology (<http://www.simile-widgets.org/exhibit/>) and allows a faceted result presentation, meaning that the user is able to choose the visualization of a given set of results using different views (currently a table and Google-maps views). Additional functionality is provided on top of the existing result set, such as filtering or sorting.

Figure 4 shows a web page built by composing the Hierarchical Search Portlet and the Faceted Result Portlet, taking advantage of portal functionality.



Figure 4 Faceted Result portlet

As it has already been discussed, while the TaToo information model is extensible, the TaToo public services interface is fixed. This is a limitation for advanced users that need intensive domain information exploitation. To overcome this limitation, the discovery service provides a method to perform SPARQL queries directly to the TaToo knowledge base. In order to embed this functionality within the TaToo Portal, the SPARQL Query Portlet has been developed. Users with knowledge on how to make SPARQL queries are free to make complex queries using this portlet and retrieve a personalized set of results according their specific information retrieval needs. Figure 5 presents the SPARQL Query Portlet.

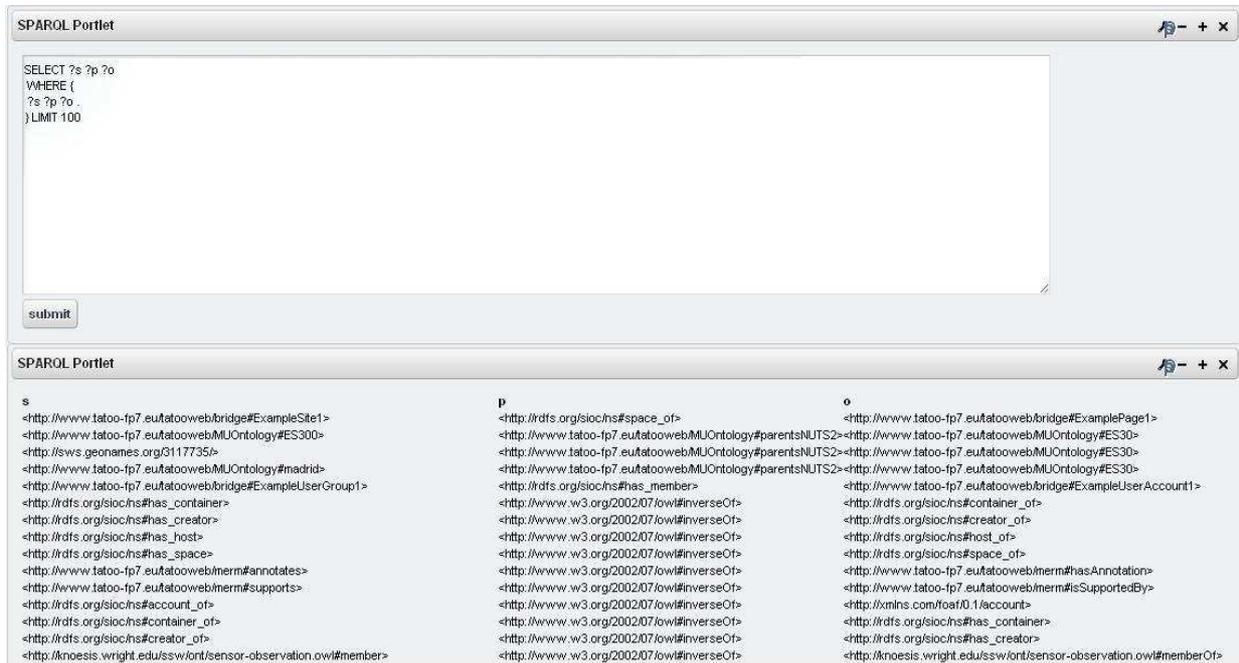


Figure 5 SPARQL Query portlet

4. Tagging

The Tagging functionality is provided in the TaToo Portal through a set of Tagging portlets. The Tagging portlets represent the Graphical User Interface access to Tagging operations provided from the TaToo core tier and exposed by the TaToo Public Service tier.

As in the case of discovery, the Tagging business logic is implemented in the TaToo core components providing RDF triples storage, filtering of visualised ontology concepts depending on user profiles, and evaluation / validation of resources and annotations (associated with resources). The knowledge base which serves the Tagging Portlets is composed by a set of domain ontologies. Within the framework these ontologies are bridged together with the MERM skeleton ontology through a set of relations which are composing the TaToo bridge ontology. Thanks to this approach the user is able to semantically annotate resources by just using concepts pertaining to its own domain, without the need of learning new terms. In some particular cases, for example while describing types of resources, if concepts are out of the scope of the user domain, terms from the bridge ontology can be displayed in the Tagging Portlet to cover the lack of information needed to properly annotate the resources.

Similarly to the discovery in TaToo, the knowledge base provided by the ontologies can evolve with the addition of new domains and users. However, this will have very small impact to the implementation of the Tagging GUIs. Moreover, the annotation workflow is also internally based on the MERM ontology, in order to have a concrete and effective discovery.

The TaToo Portal provides a 'Tagging' page that contains all tagging related portlets. At present time TaToo has implemented a general purpose Tagging Portlet, while during the second iteration of implementation a second portlet extending Tagging functionality to annotation editing and deleting is foreseen.

The TaToo Tagging portlet (see Figure 6) is composed of three different sections: ‘Resources’, ‘Tagging’ and ‘Tag Details’.

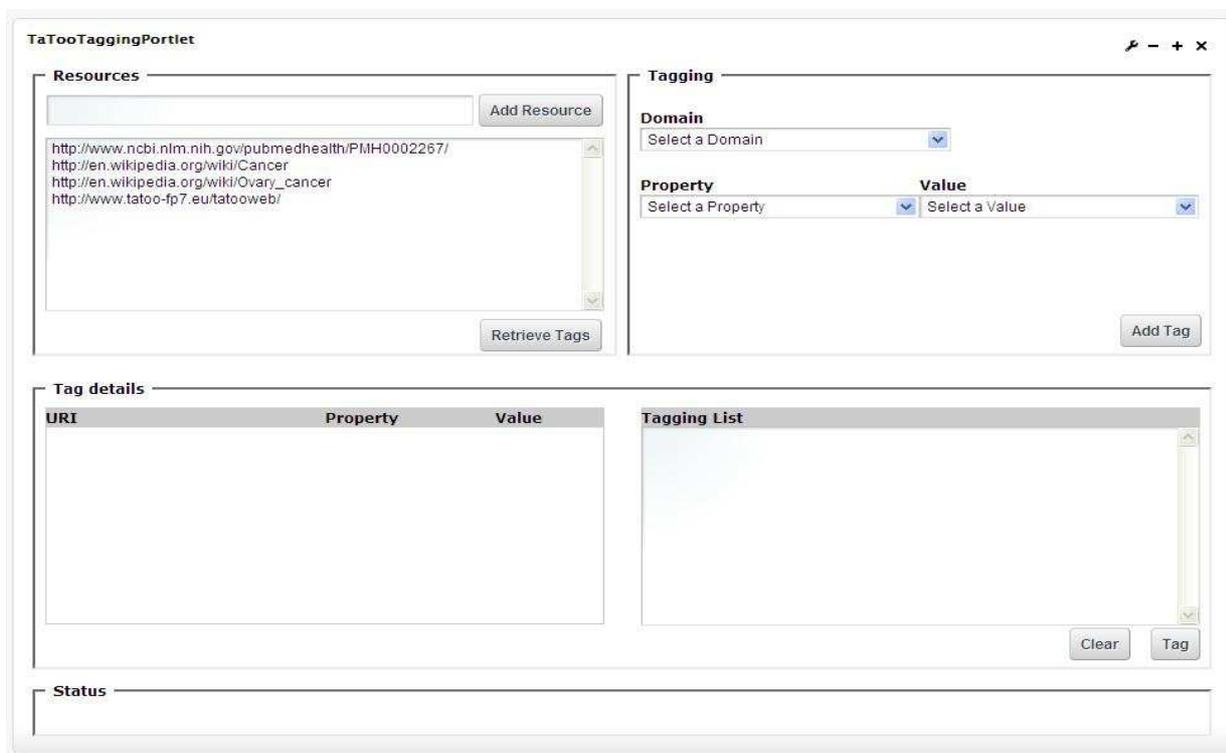


Figure 6 Tagging Portlet

In the ‘Resources’ section the user can manage their own domain resource, the user can add a new resource by providing its URI, and among the resources listed in the section, they can select a subset for which they want to either retrieve all the previously stored annotations, or they can use them to provide new annotations. The ‘Tagging’ section is used to select the instances of properties and classes that the user wants to provide to annotate the selected resources. These two information are provided as property and value respectively and are influenced by the domain selection from the user. This means that the properties will be taken from the user domain ontologies, and the values will be instances of classes from the same domain ontologies. In TaToo at the moment the selection is restricted to AIT, JRC and MU domains that are the validation scenarios, but in the future this GUI approach allows to plug-in new users domains. Once the user has selected the domain, and then the property and the value to annotate, in the ‘TagDetails’ section a list of all the annotations to be stored are presented to the user. The user can verify them and decide also to remove a few from the list. They can then finalize the annotation process by pressing the ‘Tag’ button. The selection performed by the user forms an RDF triple where the subject is the resource identified by its URI, the predicate is the property and the object is the selected value. This RDF triple is formalized in the core components tier and properly added to the TaToo knowledge base. In the same ‘Tag Details’ section, the user can also find the list of annotations retrieved from the ‘Retrieve Tags’ operation performed on the selected resource.

The TaToo Portal also presents a ‘Masaryk University’ page. In particular the Masaryk University page has been configured to provide a public environment to the Masaryk University validation scenario part-

ners, where they can deploy and take advantage of different portlets addressing their needs. In this page a new Tagging portlet has been provided to demonstrate how different GUIs can be designed and customized on top of the Public Service tier. The Masaryk University Tagging portlet represents then an example of how custom User Components can be provided to address specific requirements from different user communities (see Figure 7).

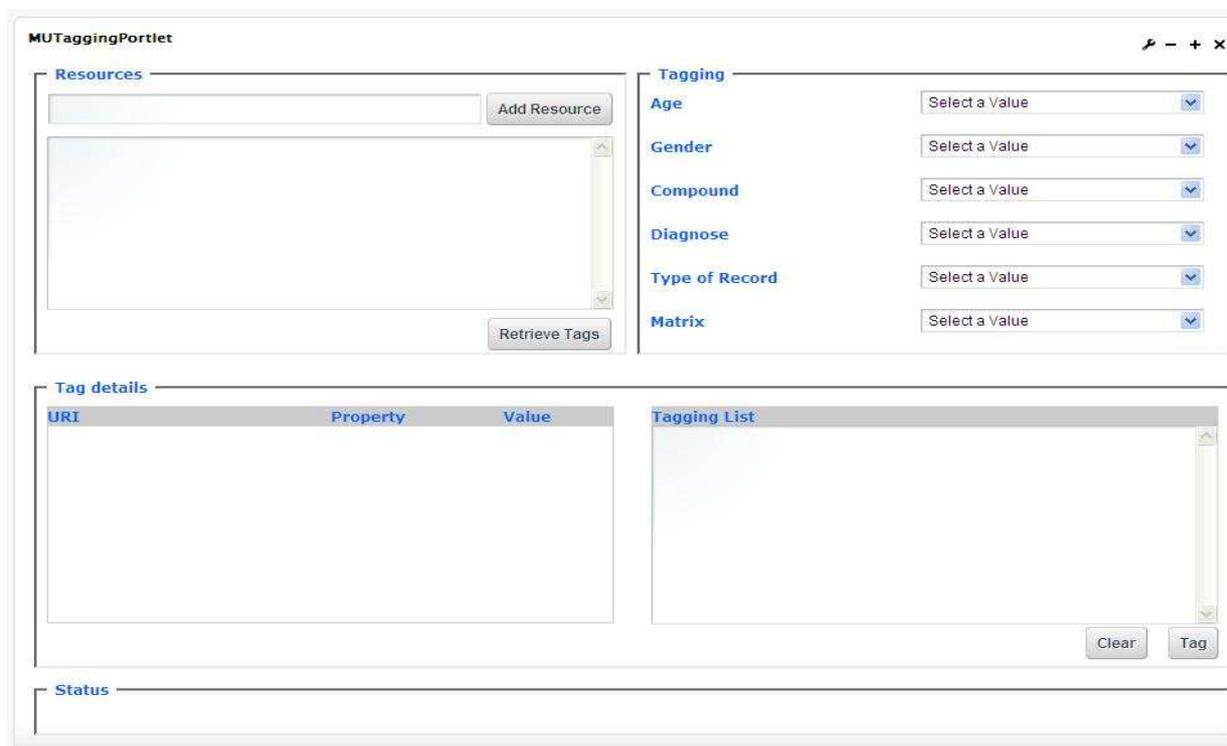


Figure 7 Masaryk University Tagging portlet

With respect to the generic Tagging portlet, the presentation of the property and value concepts has been organized to guide and help the Masaryk University community user to properly annotate resources from their environment. In particular, the property and value concepts are taken from the Masaryk University domain ontology, and the user is asked to select for a determined property, which is fixed, the correspondent class instance. For example, the “Compound” label corresponds to the property “isAbout” and the values are exclusively instances of the ontology class “Compound”, from the Masaryk University persistent organic pollutants (POP) domain ontology; “Type of Record” is bound to the property “typeOf-Record” and the values are instances of the ontology class “Epidemiological Measures”, from the Masaryk University Cancer domain ontology.

One particular use case of tagging is represented by the Evaluation and Validation. In TaToo the Evaluation and Validation consists of associating annotations to annotations. Evaluation and Validation annotations are conceptually annotations indicating the level of agreement on an already provided annotation for a resource. The development of an Evaluation and Validation portlet is at the present time ongoing and will be finalized during the second implementation phase of the TaToo project.

5. Impact and Conclusions

The aim of the presented TaToo Portal is to provide a user-friendly Graphical User Interface that delivers all the functionality provided by the TaToo Core Components. In this paper we have outlined the modularity and flexibility of the chosen approach involving a portal server and multiple portlets that can be composed in different ways by the user to set up the desired environment. Moreover, the TaToo portlets can be also customized to address specific needs, or the user can implement their own custom portlet that can be deployed and configured in the TaToo portal as well.

At the present time the TaToo project is going through its second implementation phase. The current implementation as of today has delivered already a good number of portlets pertaining to Tagging and Search & Discovery processes. However, as enlightened in this paper, the focus is now on providing to the user new appealing tools, together with new functionalities and new interfaces to enlarge the consensus of the community. A solid work is also ongoing in parallel in the TaToo business logic, especially concerning the knowledge base, where a team of ontology experts is continuously refining the TaToo ontologies, bridge, MERM and use case domains.

It is foreseen that in the third and last version of the TaToo framework, the TaToo Portal will be in its final stage and will provide a large number of portlets, addressing also evaluation and validation, geo-localization of annotated or discovered environmental resources, user context management and administration.

6. Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement nr. 247893.

Thanks to the TaToo Consortium for the support of this paper.

Bibliography

- Dihé P., Frysinger S., Güttler R., Schlobinski S., Petronzio L., Denzer R., Nešić S., Pariente Lobo T., Schimak G., Hřebíček J., Donatelli M. (2011): An Architecture for the Semantic Enhancement of Environmental Resources in Environmental Software Systems - Frameworks of eEnvironment, 9th IFIP WG 5.11 International Symposium, ISESS 2011, Brno, Czech Republic, June 27-29, 2011, Proceedings by Jiří Hřebíček, Gerald Schimak, Ralf Denzer (Eds.), IFIP Advances in Information and Communication Technology Volume 359, pp. 372-364; DOI: 10.1007/978-3-642-22285-6; www.springerlink.com
- Pariente Lobo T., Fuentes J.M., Sanguino M.A., Yurtsever S., Avellino G., Rizzoli A.E., Nešić S. (2011): A Model for Semantic Annotation of Environmental Resources: The TaToo Semantic Framework; in Environmental Software Systems - Frameworks of eEnvironment, 9th IFIP WG 5.11 International Symposium, ISESS 2011, Brno, Czech Republic, June 27-29, 2011, Proceedings by Jiří Hřebíček, Gerald Schimak, Ralf Denzer (Eds.), IFIP Advances in Information and Communication Technology Volume 359, pp.419-427; DOI: 10.1007/978-3-642-22285-6; www.springerlink.com
- Rizzoli A.E and Schimak G. et al. (2010): TaToo: Tagging environmental resources on the web by semantic annotations; Proceedings of *International Environmental Modelling and Software Society (iEMSS) 2010* International Congress on Environmental Modelling and Software Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.) <http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings>