Geo Semantic Web Communities for Rational Use of Landscape Resources

Ernesto Marcheggiani*, Michele Nucci, Giovanni Tummarello, Christian Morbidoni,

Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni
* Dipartimento di Scienze Applicate ai Sistemi Complessi - Agrur
Technical University of Marche
Via Brecce Bianche – 60131 Ancona (ITALY)
e.marcheggiani@univpm.it, mik.nucci@gmail.com,
{g.tummarello,c.morbidoni}@deit.univpm.it.

Abstract. A methodology based on Semantic Web Initiative standards and technologies in order to improve rational use of urban, rural and semi-natural global cultural landscape resources in periurban fringe expansion areas and in riconnection and ecological rebalancing green areas, is proposed. This paper aims at showing how the need of advanced cooperative annotations and information exchange can be addressed using a paradigm called “Interconnected Geo Semantic Web Communities”. Furthermore we aim at designing the bases to implement a Decision Support System to rationally manage resources which makes use of both geographical information advanced technologies such as remote sensing and Geographic Information Systems technique and semantic rules acting on ontologies. Based on the DBin platform a specific environment has been built to manage interoperable data arising from dedicated specific communities annotations and survey campaigns. The system features “collaborative annotations” and “intelligent” data structures which can be easily personalized easily.

Keywords: Ontologies; Semantic Web; GIS; DBin, Sustainable tourism; Landscape resources

1.0. Introduction

In order to adhere to the European Charter for the Sustainable Tourism, different stakeholders, belonging to thirteen municipalities of the “Comunità Montana delle Alpi Lepontine” (Cmal) in Northern Italy, cooperated using an advanced tool for Semantic Web communities, provided with geographical plug-ins which enable a real-time interaction with systems as Google Earth and Google Maps supported by specific ontologies to edit, visualize and share information.
The use cases and associated needs have been highlighted and after that the base tool for this work, DBin (G.Tummarello et al, 2006), a Semantic Web rich client application, has been focused.

DBin enables users to create and experience the Semantic Web by exchanging information structured according to Resource Description Framework (RDF) (RDF Semantics, 2004), standardized by the World Wide Web Consortium (W3C), in peer-to-peer (P2P) "topic" channels. Once sufficient information have been collected locally, rich and fast browsing of semantically structured knowledge becomes possible even offline without generating external traffic or computational load.

DBin has a number of modules to support cooperative tagging and annotations of geographical objects. Different communities of users, e.g. concerned with different kind of geographical objects, can exploit DBin to cooperate in an enriched geo-semantic space. Advanced users, e.g. cultural heritage agencies, can join multiple groups at the same time and use the collective cross domain knowledge which will be relevant not only for planning but also to allow the implementation of Decision Support Systems (DSS) for the rational use of different landscape resources.

2.0. European Charter for Sustainable Tourism in the Comunità Montana delle Alpi Lepontine (Cmal) landscape: organization, interests and needs

The case study area belongs to that large geographical region which can be identified with the northern part of Italy, considered as a possible Mediterranean Megalopolis. This territory, stretching between the Canton Ticino (Switzerland) and the Italian province of Como, presents many peculiar features. In the so-called “three lakes region” – Lugano, Piano and Como Lakes - a unique urbanistic model has been developed, since the second half of the 19th Century. It deals with that “historical villas landscape”, that occupies fringes degrading from the alpine area toward the lakes coastline. These large territorial fringes can be well divided into parallel parts with the mountainous and hilly ones characterized by rural ecosystems - here the 'alpeggio' (seasonal pasture) practice is still a productive activity that contributes to the persistence of the agro pastoral landscape. Persistence of agriculture in those areas preserves great naturalistic areas ecological functionality, acting as biodiversity reserves and thus subjected to different protection levels.

In the last decades the area has been particularly affected by severe pressure due to both a generalized uncontrolled urbanistic expansion and a relevant touristic flow coming particularly from Northern Europe. It is therefore necessary to cope with these landscape different aspects - scientific, historical, artistic and economical – from a multidisciplinary and holistic point of view which enables a complex scientific image in order to suggest integrated and effective planning actions for a rational use of natural and cultural landscape resources.

According to the Como Tourist office, the area attracts approximately 3 million persons per year, half of which from outside Italy. The area is, in general terms, welcoming and well organized according to traditional standards, but the importance to provide tourists with rich and personalized offers is quickly increasing. This requires integration and sharing of information between a number of local actors using state of
the art information technologies. It will be shown how this can be addressed with the paradigm of Interconnected Geo Semantic Web Communities.

In a wider perspective Cma1 area is only a part of a broader net of natural and historical elements that constitute the whole local cultural heritage, surrounded by a complex landscape where each stakeholder performs a number of specific tasks, most of which would benefit from cooperatively exchanging geo-spatial semantic annotations.

3.0. Geo Semantic Web Communities

The Geo Semantic Web community tool, proposed by the authors, is a specific application of DBin Semantic Web Platform. DBin enables user communities to annotate concepts of common interest and browse information structured by RDF according to ontologies for specific domains of interest.

For this study, DBin has been provided with both geo-enabled visualization and editing modules, as well as with processing capabilities to handle data structured according to geographical and geometric ontologies such as the W3C Basic Geo Ontology (W3C Semantic Web Interest Group, 2003) and RDFGeom Ontology (Chris Goad, 2004).

In this work the term “ontology” refers to a set of classes and related properties, relations and constraints defined using RDF Schema Language (W3C Consortium, 2004a) and/or Ontology Web Language (W3C Consortium, 2004b) for modelling a specific domain.

Geo Semantic Web communities are characterized by typology of knowledge that is exchanged among them. This means that by connecting to a specific community, users will only share and exchange those information which are relevant to the community itself. In DBin this is achieved by means of topic specific P2P groups based on the RDFGrowth algorithm discussed in section 4.0.

In the proposed case study, cultural heritage agencies are interested in annotating the state of cultural patrimony. At the same time parks and woods maintenance are

Figure 1. Location of the study area in Northern Italy.
performed by Forest Guards through specific actions, such as dangerous trees pruning or cutting, service buildings and routes maintenance, etc.

The proposed Rich Semantic Web client solution offers to each group a specific environment where such annotations can be edited, browsed and queried to assist decisions and management. Collaborative metadata annotations created within each community, can also point at rich media, posted on the web (e.g. pictures, documents, long texts, etc.). Users who receive annotations could then reply or further annotate it, locally, for personal use or back into public knowledge.

Figure 2. Our reference use case: groups of expert annotate and use their own domain annotations. Experts, by joining different communities can benefits of joint cross-domain knowledge.

Due to the large amount of visitors per year, CMAL area managers - e.g. advisors and majors - have broad interests which go beyond single users community. Taking care of urban area and mobility infrastructure maintenance, dwellers and visitors security, natural and cultural heritage resources conservation, etc., they are interested in performing cross-domain queries, such as unsafe touristic paths or historical building identification. Meanwhile the Local Tourist System (STL) managers are interested in attracting the largest number of visitors per year without caring about the impact of that fluxes on the local natural and cultural resources. STL aims, for examples, at informing visitors about temporary inaccessible naturalistic paths or about B&Bs lying close to naturalistic areas. Tourists are just interested in having the possibility to freely choose accommodation - hotels, B&Bs, country houses, etc. - or places in which it is possible to taste local food, experiencing a safe and positive journey.

It is interesting to underline how the use of web resources in planning holidays in an emergent and increasing market, especially among northern European cultures. Tourism is cross-sectorial, involving a wide range of issues: trade and investment policy, employment and enterprise, public-private partnerships, community and urban planning - land use planning, transportation, etc. - conservation of cultural heritage, protected areas and biodiversity, management of natural resources (water, energy, waste), safety and security, education and workforce development. There are a lot of stakeholders, with their different and sometimes opposing interests and agendas.
In this context, Semantic Web Communities can be seen from a new point of view. If two communities share resources identified by Uniform Resource Identifiers (URIs), object annotations, originally posted in a specific community, are automatically cross posted to other communities, as soon as the annotated objects is of interest to both communities.

Summarizing, actors join either one or more P2P groups to learn and exchange information about topics of interest, thus collecting an heterogeneous amount of structured knowledge expressed in RDF.

In this context the issue becomes how to allow users to interact with RDF data in a natural way. This is a strongly domain-dependent task, as each community deals with different kinds of information and has different needs in terms of browsing and editing capabilities.

The proposed solution is to provide a way for a "group leader" to define and make available to users, a set of domain-specific interaction environments called Brainlets, described in section 5.0.

New Geo Semantic Web Communities can be started at will, with a relative little technological effort, by defining P2P groups for specific topics and “publishing” related Brainlets.

4.0. The RDFGrowth P2P engine: high level overview

RDFGrowth algorithm is used by DBin clients to collect and distribute RDF data within groups of interest. The algorithm is presented and discussed in G.Tummarello et al, 2004; a high-level overview is given here.

Previous P2P Semantic Web applications, such as Wolfgang Nejdl, Boris Wolf 2002, Paul Alexandru Chirita 2004, Min Cai, Martin Frank 2004 and Wolfgang Nejdl 2003, have explored interactions among groups of trusted and committed peers. In such systems, peers rely on each other to forward query requests, collect and return results. On the contrary, the real world scenario of peers, where cooperation is relatively frail, is here considered.

In RDFGrowth peers are certainly expected to provide some external service, but commitment is minimal and in a “best effort” fashion: no commitment in terms of complex or time-consuming operations, such as query routing, collecting and merging, is required from peers.

4.1. Overview and main features

RDFGrowth algorithm enables the creation of P2P groups around topics of interest. When an user joins a specific P2P group, a software agent, based on RDFGrowth algorithm, begins to collect and share only information about the group’s topic. Topic definition is given by the Group URI Exposing Definition (GUED). GUED is an operator which, applied to a RDF knowledge base, returns all and only those resources conforming to a set of semantic constraints (e.g. “hotels and restaurants located within a naturalistic area”). Only the resources extracted by a GUED operator will be shared with peers of a same P2P group of interest.
Peers in the same group exchange among themselves pieces of information extracted locally at each peer by a GUED operator, so that, after a transitory period, each of them will exactly have the same knowledge about the group’s topic and each of them will have to store locally the knowledge obtained from the P2P communication channel.

The RDFGrowth approach is particularly suitable in every scenario in which users have necessity to work with off-line data and then, after having processed it, they will have to exchange results within a group of interested people. For instance, this approach could be applied to a geo-annotation scenario in which every user (every peer) could work off-line on mobility devices far from a connectivity point and only after having annotated some locations (e.g. “tourist facilities in a naturalistic area”) they will exchange the resultant information among them.

P2P groups and related GUED operators can be defined by editing an XML configuration file. In particular, a GUED operator can be implemented as a set of Semantic Web queries based on SeRQL query language (Jeen Broekstra, Arjohn Kampman, 2004) which selects and extracts information of interest from an RDF knowledge base. Example 1 shows a possible GUED configuration for a group interested on tourist facilities within naturalistic areas in a specific municipality.

Example 1. The Tourist Facilities group configuration.

```xml
<group name='TouristFacilities'>
  <gued query='SELECT X FROM {X} rdf:type {http://dbin.org/buildings#touristFacilities}; location:locatedIn {Y} rdf:type {http://dbin.org/area#naturalisticArea} WHERE Y = "http://dbin.org/municipality#13047" />
  <default_brainlet name="TouristFacilities" uri=http://dbin/brainlets/touristFacilities />
</group>
```

4.2. Identities and authorship of information

In a potentially large and unregulated P2P community it is important to have information about “who said what”, in particular which user is the author of a particular annotation received from the network. To enable this, in DBin a methodology based on the Minimum Self Contained Graph (MSG) theory defined in [13] is used. Such methodology allows each piece of information inserted by users, to be digitally signed at a fine granularity level and in an efficient way. It also assures that authorship information will remain within metadata when they will be exchanged by P2P communication channel. Details about digital signature process are outside the scope of this discussion and can be found in G.Tummarello et al, 2005; only a high-level overview of the basic procedure is given here.

When started up for the first time, DBin clients require the users to provide a valid URI or a nickname (used to automatically generate a valid URI to indentify) which will act as an identifier for the user itself. Then a public and a private keys are gener-
ated; the private key is stored locally, while the public key is uploaded to a public server. Every time a user will insert an annotation into the system, the user identifier as well as the URL of the public key will be added to the annotation itself which will be also signed using his/her private key. In this way, after having received a piece of information from a P2P group, DBin clients are able to retrieve the associated public key to identify the author of the annotation, without caring about the provenance of the information itself.

Once the authorship of each annotation can be derived, a variety of local filtering rules can be applied at will. For example, users can build a local policy in order to hide annotations from untrusted authors.

5.0. Brainlets for Semantic Web Communities

Defining Semantic Web communities and working with data coming from specific domains of interest requires dedicated software environments and tools such as data visualization systems, specific data editors, etc. Brainlets can be thought as “configuration packages” setting DBin to operate on a specific domain, providing also needed tools. Brainlets can be viewed by users as full “domain applications running inside DBin”. By Brainlets it is possible to define DBin settings for:

- general UI layout, defining also involved UI components and interaction among them;
- ontologies to be use for RDF data and for annotations in a specific domain;
- template for specific domain “annotations”;
- templates for readily available “precooked” domain queries;
- templates for wizards to insert new specific domain resources;
- suggested trust models and information filtering rules (e.g. identities of “founding members” or authorities).

Since DBin is base on the Eclipse Rich Client Platform technology, every Brainlet, which technically is an RCP plug-in, can be installed into DBin in a simple way.

New Brainlets can be created by editing an XML configuration file: no programming skills are required. Other than a basic knowledge of the XML, every task involved in the process of creating a new Brainlet are related to “knowledge engineering” aspects such as: selecting appropriate ontologies to describe a specific domain of interest or making queries to select a set of interesting data.

This section describes the main elements involved in creating Brainlets related to a landscape resources planning scenario like that focused on this case study.

5.1. Ontologies and Basic RDF Knowledge Base

In creating a new Brainlet an important step is the choice of the ontologies which formally describe the concepts related to the domain of interest. Once ontologies have been selected, related RDFS/OWL files can be included into the Brainlets itself although they could be placed on the Web.
In the case study, existing ontologies have been used to foster information reuse and interoperability. In particular, the W3C Basic Geo Ontology has been used to represent basic geographic concepts such as points with latitude, longitude and altitude from the WGS84 reference datum specification. In addition to this ontology, RDFGeom ontology has been used to represent more advanced geometric concepts such as: areas, transformations, curves, segments, polylines, boxes, circles, etc. Example 2 shows the use of W3C Basic Geo Ontology to represent the location of a hotel and the example 3 shows the use of RDFGeom Ontology to represent a simple polyline.

```
<buildings:hotel rdf:resources="http://dbin.org/hotel#1">
  <geo:lat>55.701</geo:lat>
  <geo:long>12.552</geo:long>
</buildings:hotel>
```

**Example 2. representation of the location of a hotel by RDF based on the W3C Basic Geo ontology.**

```
<geom2d:Polyline rdf:resource="http://dbin.org/pline#1">
  <geom2d:points>
    <rdf:Seq>
      <rdf:li>
        <geom2d:Point>
          <geom:x>-123.835414</geom:x>
          <geom:y>46.191617</geom:y>
        </geom2d:Point>
      </rdf:li>
      <rdf:li>
        <geom2d:Point>
          <geom:x>-123.821983</geom:x>
          <geom:y>46.189905</geom:y>
        </geom2d:Point>
      </rdf:li>
      ...
    </rdf:Seq>
  </geom2d:points>
</geom2d:Polyline>
```

**Example 3. representation of a simple polyline by RDF based on the RDFGeom and RDFGeom2D ontologies.**

5.2. Navigation of RDF Resources

The way resources are presented and browsed is crucial to interface usability and effectiveness in finding relevant information. RDF data structure representation is based on a graph model. Today many graph-base visualizers exist but the problems related to the usability and usefulness of this approach are well-known: large graphs with a lot of nodes are difficult to render and navigate. Normal users are familiar with a folder-like structure so, in DBin, it has been decided to implement a Resources Navigator based on a flexible and dynamic tree structure. This approach fits better with respect to the number of resources. Resource Navigator has a completely configurable semantic multiple branches folder tree structure, each one is able to contain specific resources with respect to criteria specified by the Brainlet author.

In this case study, different Navigators have been defined. The Tourist Facilities Navigator, shown in figure 3 has been configured to have more branches, each one
can organize different kind of tourist facilities; Municipalities Navigator is shown in figure 4.

5.3. “Precooked queries”

While creating a Brainlet the author can identify for the domain of interest some queries which can be frequently used to fulfil relevant use cases. The author of the Brainlet can define these queries, called in DBin “pre-cooked queries”, within Brainlet XML configurations file.

The pre-cooked queries will be ready-to-use by the end-users of Brainlets, who will only have to “fill in the blank boxes” on the dedicate window (figure 5).

For the case study focused here, such a query could be: “find all touristic facilities in the naturalistic area “X” and areas of municipalities involved”. The results of precooked queries can be viewed in a tabular way or can be sent to a dedicated visualizer (e.g. visualization plug-ins based on Google Map/Google Earth) by the “semantic clipboard” which is a DBin clipboard for heterogeneous data semantically structured.

5.4. URI Wizards

In the Semantic Web, each resource is identified by a URI. To foster knowledge interoperability within communities, a methodology is needed to avoid different users to choose different URLs to identify the same concept. This can be achieved by defining procedures to assist users in assigning identifiers to new instances. These procedures can be defined by the author of a Brainlet and can be encoded into the so called
“URI wizards”. By XML configuration it is possible to define customized URI Wizards to assign meaningful URIs to each specific class of resources.

In the case of geographic objects, a simple “URI wizards” can be configured to obtain URIs from their natural aspects. For example an identifier for a touristic facility can be derived from its address as well as an identifier for a naturalistic area can be obtained from the coordinates (latitude and longitude) of its centroid.

Solutions for minting identifiers on Semantic Web are still in their infancy. The URI wizard approach offers flexibility to accommodate future methods as they will be made available and/or reach popularity.

5.5. Custom domain dependent annotation templates

In DBin, end-users can add new properties to a resource of interest. Brainlets use ontologies to assist users in this operation, suggesting which properties can be associated with a resource. However for some specific domains of interest, the author of a Brainlet could have the necessity to define annotations with a complex structure. This type of annotation, which in DBin is called “complex annotations”, can be defined using an ad-hoc ontology.

When a user selects a resource to add it a new complex annotation, DBin is able to determine which type of annotations can be applied to the specified resource and provides a dynamic wizard.

For example, in the Brainlet developed for the use-case considered here some, geographic complexes annotations have been defined. Figure 6 shows one of the defined geographic annotations and related wizard which permits to users to fill geo-spatial information such as latitude, longitude and altitude of a resource (e.g. a touristic facilities).

![Figure 6. An example of a geographic complex annotation and related wizard used to insert geo-spatial information.](image)

5.6. Brainlets and Geographics Plug-ins

Brainlets can also use additional plug-ins to address domain specific needs. For the considered case study specific geographic plug-ins, which permit a real-time interaction with Google Maps an Google Earth, have been used. The plug-in based on Google Maps has been used to show to end-users geographic information like the lo-
cation of a building as well as it has also been used as “annotation tool” by which an end-user can add new geo-spatial annotations only selecting a resource and pointing in its real location on the map. The plug-in which enables real-time interaction with Google Earth has been used to show complex geographic information (such as regions, paths, etc.) or, for example, to show the results of a complex pre-cooked semantic query.

In figure 7 the Google Maps-based module is showing the geo-location of a selected resource (the location of a tourist facility) while in figure 8 it is possible to see Google Earth showing the result of a cross-domain pre-cooked query (in this specific case: “show all tourist facilities within a naturalistic area and the related involved municipality).

![Figure 7. Google Maps-based module showing the location of a resource basing on geo-spatial information structured by RDF.](image)

![Figure 8. Real-time interaction with Google Earth which shows the results of a cross domain pre-cooked query.](image)

### 5.7. Social Model

Brainlets, by providing an aggregation medium for ontologies, users and data representation structures, are therefore good catalyst of overall semantic interoperability process. As users gather around popular Brainlets for their topic of choice, the respective suggested ontologies and data representation practice will form an increasingly important reality. If someone decided to create a new Brainlet or Semantic Web ap-
plication in general which could target the same user group as the said popular Brainlet, there would be an evident incentive in using compatible data structures and ontologies.

6.0. Related Works and Conclusions

Tourism is one of the world’s largest economic sectors. Sustainable tourism development can be a powerful tool for economic growth, natural and cultural resources conservation. While tourism represents an important development opportunity for many countries and communities, it can also have very negative impacts, such as disrupting and harming the socio-cultural authenticity of local communities, and threatening natural and cultural heritage.

Communication has a huge role in supporting sustainable tourism development. Communication can create and facilitate a system that allows stakeholders to exchange opinions and arrive in a rational planning at consensual solutions. Effective use of communication tools can also link products to markets, and can contribute to visitors’ safe and positive experiences (Communication and Sustainable tourism, 2006).

In general, the advantages that ontologies provide for geographic information processing include the enhancement of communication, systems engineering, and interoperability (F. Reitsma, K. Hiaramatsu, 2006). Many approaches to the visualization of semantically annotated data have been proposed in literature. RDFGravity exploits graphs as a visualization tool of RDF triples (see Sunil Goyal, Rupert Westenthaler). In E. Pietriga 2002, an environment for both RDF browsing and authoring is discussed. Welkin (SIMILE a) is an RDF visualizer based on elliptical zooming of the connections among resource, while Longwell (SIMILE b), implements the idea of faceted browsing of RDF data. MIT Haystack (Quan, Dennis and Karger, 2004) is a tool for development of Semantic Web applications, also focuses on interface organization (layout and functionalities).

While pro and cons can be argued for each specific approach, it is clear that user interface issues are complex with no clear single solution. DBin responds to this by the Brainlet UI interface, which enables a topic specific “mash-up” of different visualization paradigms.

DBin stands out as an end user application which provides Semantic Web capabilities and a plug-in based open source, rich client approach.

The combination of these two, effectively enables efficient extension and integration into domain specific tools and ontology based browsing, querying and searching information. The addition of proper geographical annotations handling components enables “Geo Semantic Web Communities”, which can be quickly started by domain experts.

Especially when compared to full featured Web GIS (as the one in Mahesh Rao et al, 2006) or specific software, the support for geo-spatial use cases is today still limited to relatively simple objects, in particular to those that have a well specified location, that is, those that can be approximated with a point.

The purpose of this work, however, it is not that of matching the state of the art specific tools. The potential of Semantic Web based annotation communities seems
unprecedented in terms of flexibility and for the ease of integration of information across communities. This is what the project seeks to explore and turns to be particular important when information is useful to support decision in whole landscape resources participatory planning (Marcheggiani et al. 2007).

It is clear how the full potential of Geo Semantic Web Communities and of the paradigm in general is to be explored and validated. This is however the case for the whole Semantic Web initiative and with respect to this, the contribution of the DBin platform and the applications that have been shown here is to represent, arguably, the most tangible instantiation of such technologies available for actual use today.

DBin is distributed under the Gnu Public License and the default distributions includes all the geographic visualization and editing plug-ins that have been used for the screen-shots. Further documentation and compiled executables can be downloaded at http://dbin.org, where it is also possible to find a few minutes screen demo which shows the whole DBin paradigm and platform.

References


