

SWAD-Europe extra deliverable 3.20: Report on developer workshop 9 - Geospatial information on the Semantic Web.

Project name:

Semantic Web Advanced Development for Europe (SWAD-Europe)

Project Number:

IST-2001-34732

Workpackage name:

3 Dissemination and Implementation

Workpackage description:

<http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-3>

Deliverable title:

Extra deliverable 3.20. Developer Workshop 9 Report

URI:

http://www.w3.org/2001/sw/Europe/reports/dev_workshop_report_9

Author:

Charles McCathieNevile

Abstract:

This report summarises the ninth developer workshop, held in Budapest, Hungary on 4 and 5 October 2004. The workshop explored how to use the Semantic Web to record geospatial information

STATUS:

Completed report The first version of this report was published 10 October 2004. Final version: \$Date: 2004/10/29 15:29:49 \$.

Please send any comments on this report to the general RDF Interest mailing list www-rdf-interest or to the geowanking list, which focuses on the topic of the workshop.

Contents

[Executive summary](#)

- [Introduction](#)
- [Motivation](#)
- [The Workshop](#)
- [Outcomes](#)

[Appendix A. Projects and Tools](#)

Executive summary

This workshop brought together developers interested in the use of the semantic web for geospatial information.

It was organised in conjunction with Progos.hu, a semantic web development company in Budapest, Hungary, and took place at the offices of the W3C in Hungary, which are hosted by Sztaki, the national institute for computer science.

The workshop had the following outcomes:

- Development of some ways to describe geospatial relations
- Introducing a range of issues to the discussion of geospatial information.
- Building on, and further motivating, the discussion of this topic in international fora
- Motivating the further development of tools for providing geospatial information

1 Introduction

This report is part of the SWAD-Europe project Work package 3: Dissemination and Implementation. It describes the developer workshop "Geospatial information on the Semantic Web", held in Budapest, Hungary, 4-5 October 2004

The principal objectives of the workshop were

1. Bring together developers working on the use of geographic data in the semantic web.
2. Provide a brief survey of available tools and vocabularies.
3. Identify areas where further development would be useful.

2 Background

Geospatial information has been an issue at the fringes of the SWAD-E project, and indeed to some extent at the fringes of the Semantic Web for some years. Along with event information it represents one of the very common areas of human knowledge that are very difficult to represent accurately in machine-readable terms.

The geowanking mailing list was set up early in the life of SWAD-Europe (but independently of the project) as a community forum. It has had a steady participation from a number of developers working in the area. However it has proven difficult for existing work to gain significant traction, largely it seems because of concerns over the interoperability of information itself - in other words, because the difficulties of modeling spatial information have not been easily overcome.

3 The workshop

Attendance - The bulk of the workshop participants represented commercial developers, and the majority of people present were Hungarian.

4. Outcomes

The workshop developed two basic tools for describing geospatial information. It also discussed in some depth the issues that surround description of geospatial information. It is hoped that this discussion will form the basis of an explanation of these issues providing a clear guide to the minimum amount of information necessary to usefully describe geospatial information in the semantic web.

Use cases - Traffic flow

Being able to determine a large number of time-stamped locations for people is one way of doing traffic analyses. Even without explicit information about when people are stopping or slowing down on their own initiative, it can be used to determine some basic information about the maximum speeds being attained, and some information about where people are choosing to go.

Finding accuracy

People generally describe places in terms of a named area, with some kind of bounded space identified by the name. Although these names are not unique in many cases, people are capable of disambiguating in most contexts. By contrast most geospatial information used on the web today is described in terms of particular coordinate locations, often with the accuracy of that information unclear. Describing the fact that a point in the middle of France is within France is fairly straightforward. But describing whether a point near the border really is within the border, or just appears to be because the definition of the border that is available approximates without enough accuracy, is difficult.

There is the "GeoOnion" vocabulary for describing concentric sizes of "points", and there are several vocabularies used describing that something is "near" something else. But the first of these implies an area described by a radius which often does not correspond to the actual shape of places, and the second does not provide much notion of how near "near" means.

Meeting people

Related places - in and out - One of the simplest descriptions of places that has gained wide success is various aspects of the concept "near". While these are extremely imprecise, it seems that various usage contexts, such as "a person is near an airport", or a restaurant is near a given geographical point, contain enough information to infer something useful about the accuracy (or otherwise) of the relation.

Exact relations are harder to define. The workshop did decide to produce a small vocabulary for describing the exact relation that something is geographically contained by something else, and the logical opposite of that, that something is not at all contained by something else. In order to cope with the Semantic Web's nature, a third property was developed to describe things which are partially in and partially not in something else.

Points

Since many places are described as a point, with varying degrees of accuracy, it is helpful to determine that some points are fully contained within others. The use of coordinates accurate to within 2 meters provides point locations different to the use of coordinates accurate to within half a mile. Yet objects as large as airports are often described as having a geographic location which is a point, with the point specified being less than a centimetre across. In the context of an area larger than 100 metres, it is helpful to note that a given point is in the area, rather than asserting that it is an accurate description of the area's location.

Irregular shapes

Very few places have the ellipsoid shape that is actually defined by a particular latitude and longitude coordinate pair. When an approximate location is described for a rectangular building, it can be helpful to describe smaller ellipsoid areas that fall entirely within the building - for example at the corners. Likewise it is possible to determine where within a particular ellipsoid a building is by describing where it is not - again through the use of highly accurate latitude and longitude coordinates.

Named places

Most place descriptions used by people are in fact names, rather than geometric projections of geospatial location. Being able to describe that various named places are inside each other is important to establishing some measure of interoperability. This is a technique often implicit in postal addresses, which describe zones of containment from the level of a country through the area covered by a postal code down to the detail of a particular apartment or room. In order to use this information in the Semantic Web it is important that we have a way to describe these relations explicitly.

An example that describes the fact that two GPS coordinate locations are within the meeting room where the workshop was held

```
<geo:SpatialThing rdf:ID="A0">
  <dc:title>W3C Meeting room</dc:title>
  <geox:in>
    <geo:SpatialThing>
      <dc:title>Sztaki</dc:title>
    </geo:SpatialThing>
  </geox:in>
</geo:SpatialThing>

<geo:SpatialThing>
  <geox:in rdf:resource='#A0' />
  <geo:long>47.477836833333</geo:long>
  <geo:lat>19.051771</geo:lat>
</geo:SpatialThing>

<geo:SpatialThing>
  <geox:in rdf:resource='#A0' />
  <geo:long>47.47783683</geo:long>
  <geo:lat>19.05177121</geo:lat>
</geo:SpatialThing>
```

France is not in New Zealand

```
<geo:SpatialThing>
  <dc:title xml:lang="fr">France</dc:title>
  <geox:notIn>
    <geo:SpatialThing>
      <dc:title xml:lang="mi">Aotearoa</dc:title>
    </geo:SpatialThing>
  </geox:notIn>
</geo:SpatialThing>
```

Some parts of Spain are outside the European duty free area

```
<geo:SpatialThing>
  <dc:title xml:lang="fr">Espagne</dc:title>
```

```

<dc:description xml:lang="en">All of Spain, including the Canaries etc</dc:description>
<geox:innish>
  <geo:SpatialThing>
    <dc:title>The European common duty zone</dc:title>
  </geo:SpatialThing>
</geox:innish>
</geo:SpatialThing>

```

A more complete version of [these three examples](#) is available.

Defining regions - BorderSegment and rdfgeom - the relation between the two, the implications for tools of a fuzzy border.

Borders are not neat and tidy, and it's useful to know roughly how accurate they are. The endpoints are not clear to start with, and then the area between them may be a rough approximation, useful for a given use case. To find out if it is useful for another use case you need to know more about it.

There is a [borders vocabulary](#) that provides for a border segment which is essentially the space bounded by two points and all possible straight lines between them, with an optional "fuzziness" value (whose datatype is not clear) to increase or decrease the area of uncertainty. A BorderSegment is partially in each of the things whose common boundary it delimits.

Addresses - Postal addresses are an example of a common scheme for identifying geographical places. They can provide an address as small as a little mail box in a post office, or via *Poste Restante* add additional semantics - effectively similar to those of the nearest Airport than the traditional understanding of the area described by "a street address". Similarly, the valid Australian postal address *Joe Citizen*

Yuendumu via Alice Springs 0800

Localises a person for the purpose of the use case, but not in the same way as

Dra Juanita Ma. Lopez

123 c/- Alcaldo Hombre 4o D

12345 Lugar Conocido

Repblica Argentina

A simple [ontology](#) for postal addresses, designed explicitly to describe Hungarian postal addresses, and an [example](#) Hungarian address marked up using the ontology were developed during the workshop as demonstrations of this approach. More work in this area would be useful.

Work still needed - [Converting BorderSegment to simpler markup for generating maps](#)

The BorderSegment method provides a powerful way of describing regions. But although the encoding is very simple, it is also verbose - the [fuzzyTriangle](#) example describes a region bordered by 3 segments, and is half a page long. By comparison, the equivalent rdfgeom (which copies the SVG model) is

```
<region rdf:value="M 12 34 34 34 12 12 12 34"/>
```

Describing the relation between the two, and how to use the relation to build real SVG maps would be a useful piece of work. In particular, a solution should take account of how to ensure that the border segments are active, and can be used either to let the user know that a particular point is not clearly on one or other side of the border, or to fetch more precise information that can be used to determine this.

[Simplification and clarification of border segments](#)

It may be possible to simplify the BorderSegment approach, just using a sequence of "points" and describing the accuracy around each point.

The current definition assumes a clockwise motion to describe things inside the border, and an anti-clockwise sequence to describe regions which are within the border of the area described but do not form part of it. Is there another rule that can be used?

[Address schema](#)

Addresses vary significantly around the world, but they are a very useful way of describing locations, with a large amount of information being keyed to addresses (where people reside, work, or meet, the boundaries described by an address, as well as a well-defined transport protocol identifier).

During the workshop some work was done to describe Hungarian Postal addresses, identifying the parts that are common to each. There are XML and other vocabularies describing addresses, but as far as we are aware there is no RDF vocabulary that takes into account both the wide range of regional variations and the fact that some components are common across a number of different types of address. The initial [design](#) for an address vocabulary sketched in the workshop should be tested and extended to cover other types of postal addressing scheme.

Appendix A Projects and Tools

There are a great many projects and tools in this area, and those listed below are only those explicitly touched on during the workshop.

DAML.org airport lookup

A service that returns RDF data about an airport, including its location, given its IATA or ICAO code - for example <http://www.daml.org/cgi-bin/airport?BUD>

☞ Foaf people map

A map that provides a visual representation of how many people have recorded information about where they are, using the FOAF vocabulary.

☞ GeoX

A small vocabulary of geospatial terms, including "in", "not in", and "partially in", which resulted from the workshop, along with others designed to extend the wgs84_pos vocabulary

☞ Carte zones monde

An interactive map that allows the user to generate RDF describing a particular point on a map. This was further developed as a result of the workshop.

☞ Nearest Airport

A simple tool from Morten Frederiksen that gives information about the nearest airport given a latitude and longitude, or returns information about an airport given its IATA code.

☞ PhotoRSS

A project working on attaching location data to photographs, for RSS feeds or Webpages.

☞ Space.frot

A collection of work by Jo Walsh on representing geographic information in RDF

☞ wgs84_pos

A vocabulary widely used to describe points in space. It defines two terms for parts of space - a spatial thing, and a point, along with three basic terms for describing their location: Latitude, Longitude and altitude. The first two are defined in terms of latitude and longitude coordinates as defined by wgs84, while the latter is essentially currently undefined.

☞ world map

A map that shows countries of the world as a highlight. As a result of the workshop, features of this map are being used to dramatically improve the functionality of carte zones monde.