D2.5.2: Demonstrator of the Natural Language Discovery and Query Interface

Title	D 2.5.2: Demonstrator of the Natural Language Discovery and Query Interface
Creator	Claudia Cialone, Kristin Stock
Creation date	20/06/2011
Date of last revision	30/06/2011
Subject	multilingual, Natural Semantic Metalanguage, semantics, geospatial, WPS
Status	☐ Draft ☒ Final
Publisher	EuroGEOSS
Type	Text
Description	This document contains the user instructions for the natural language discovery and query interface demonstrator that is Deliverable 2.5.2. It also includes screen dumps and links to the demonstrator itself and a video demonstration.
Contributor	Claudia Cialone, Kristin Stock (CGS, University of Nottingham)
Format	Word document.
Source	
Rights	☐ Restricted ☒ Public
Identifier	EuroGEOSS_D2_5_2.doc
Language	En
Relation	WP2
Coverage	Not applicable

These are Dublin Core metadata elements. See for more details and examples http://www.dublincore.org/

TABLE OF CONTENTS

FIGURES

ACRONYMS AND ABBREVIATIONS

Abbreviation	Name
IOC	Initial Operating Capacity
GUI	Graphical User Interface
AOC	Advanced Operating Capacity
CNR	National Research Centre
JRC	Joint Research Centre
SBA	Societal Benefit Area
WPS	Web Processing Service
CSW	Catalogue Service Web
WP	Work Package
WPS	Web Processing Service
NSM	Natural Semantic Meta-language
EEA	European Environmental Society
GEO	Group on Earth Observations
GEOSS	Global Earth Observation System of Systems
XMLHTTP	Extensible Markup Language Hypertext Transfer Protocol
OGC	Open Geospatial Consortium
GIS	Geographic Information Systems
SKOS	Simple Knowledge Organization System
RDF	Resource Descriptor Framework
OWL	Ontology Web Language

1 INTRODUCTION

The EuroGEOSS project involves, inter alia, the creation of a natural language query interface for the discovery of geographic resources. We have provided a graphic user interface that represents a first step towards allowing users to perform natural language queries in a range of different languages. At this stage, the interface allows natural language spatial queries to be performed in English, but has been designed so that other languages can be easily added in the future; and later research will widen the scope of queries to include temporality. Behind the interface is an OGC compliant Web Processing Service (WPS), that developers may wish to access directly. Details of this WPS can be viewed through the GetCapabilities and Describe Process methods on the WPS:

- http://128.243.249.134:8080/wps/WebProcessingService?REQUEST=GetCapabilities&SERVICE=WPS and
- http://128.243.249.134:8080/wps/WebProcessingService?REQUEST=DescribeProcess&SERVICE=WPS&version=1.0.0&identifier=uk.ac.nottingham.cgs.TestProcess.

The aim of the interface is to allow users to combine:

- ontology concepts (and semantically similar concepts) to specify geographic features of interest (for example, forests, drought-affected areas);
- natural language terms to formulate a spatial query between the geographic features of interest and
- optionally, geographic coordinates to narrow down the query to a particular geographical area.

Below is a brief presentation of the interface and a description of how to interact with it. A video demonstration can be viewed at: http://128.243.249.134:8080/NLInterface.2.1.1-beta/MNL_WPS_Demo_tutorial.html

2 PRESENTATION

2.1 What Does the Interface Look Like?

The interface can be viewed at http://128.243.249.134:8080/NLInterface.2.1.1-beta/ (see Figure 1 below). The interface requires the Mozilla Firefox v5.0 browser. It may also work with other browsers, but it has only been tested with Mozilla Firefox v5.0.

The interface includes:

- a list of available ontologies (at the moment only the GEOSS SBAs and GEMET but more specific ontologies (or thesauri) will be added later, for example, the drought vocabulary) and a box that allows selection of concepts and the language to be used;
- a blank box where a natural language query can be formulated (the NSM box in Figure 1 below);
- a set of semantic primitives (Natural Semantic Metalanguage or NSM) to express spatial relations;

- a map to allow users to specify a geographical area of interest and
- a help button, a send button, and the server address of the application..

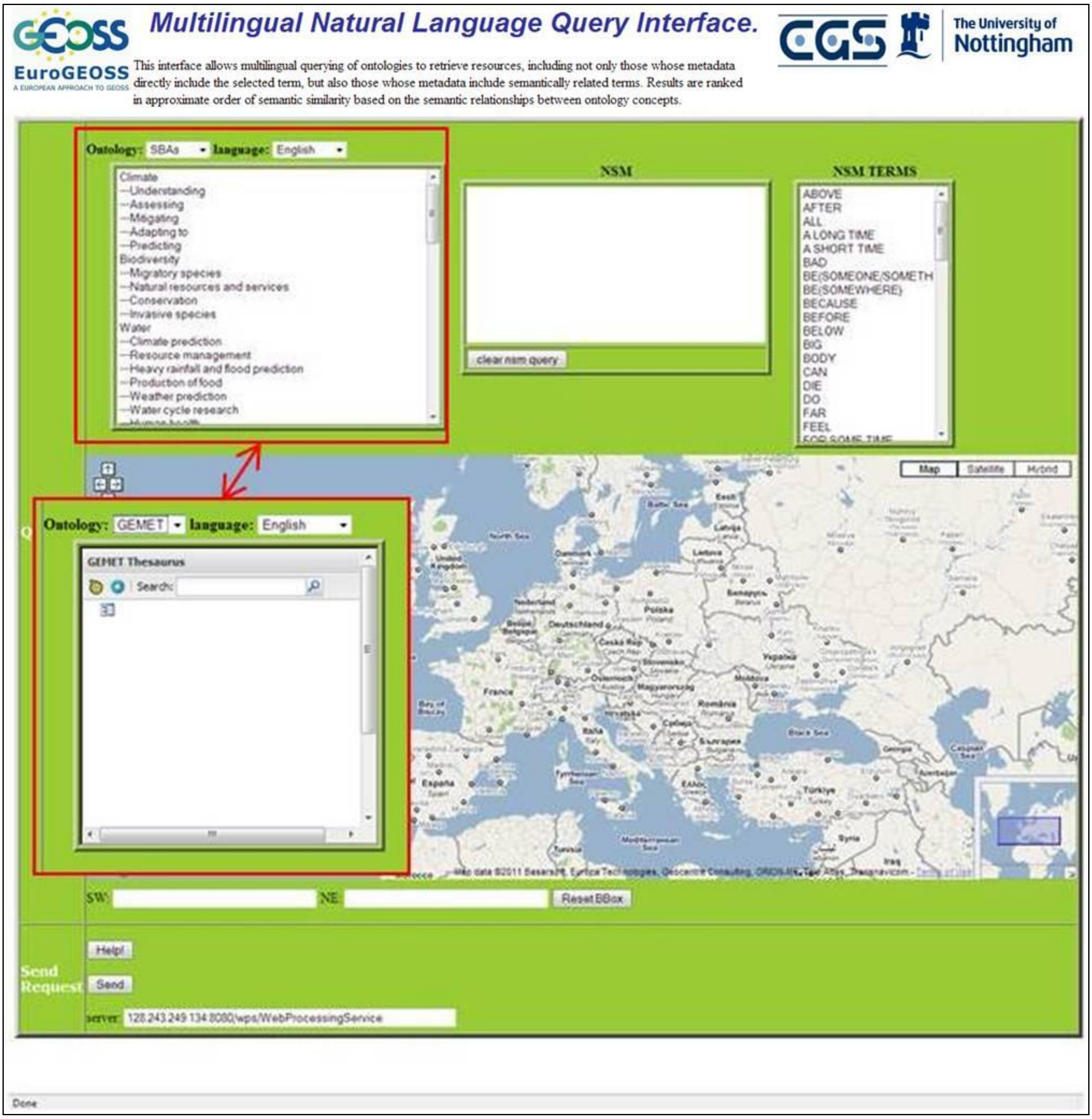

Figure 1: Natural Language Query Service screenshot

2.2 About the Ontologies: GEOSS SBAs and GEMET

GEMET[1] is the Multilingual Environmental Thesaurus containing more than 6,000 descriptors and covering a wide range of environmental issues. It was originally developed by the European Topic Centre on Catalogue of Data Sources (ETC/CDS) under contract to the European Environment Agency (EEA) based in Copenhagen and its Network (Eionet).

The Societal Benefit Areas (SBAs)[2] are a set of 9 categories and 58 subcategories established by the Group on Earth Observation (GEO) for the GEOSS project, as a thematic endpoint for Earth observation and monitoring.

In the following example, GEMET is shown (bottommost red square on the left) with a particular interface but this can be toggled with the SBAs (a simple list of concepts in the other red square on the right) as shown as an alternative.

2.3 What Does the Interface Do?

The natural language interface allows the user to select two ontology concepts of interest from GEMET and/or the GEOSS SBAs categories and subcategories in any supported language by browsing over them. Resources are in fact retrieved thanks to the interrogation of these two main ontologies adopted as semantic support for the project: GEMET and the GEOSS SBAs, provided the resource's metadata has been previously tagged with some of the concepts from the ontology shown.

The query interface allows the user to refine the search by formulating a spatial query in NSM (in the form of a spatial relation between two ontology concepts) that filters the results and by identifying a bounding box to specify a geographical area of interest.

The whole interface is intended to operate within the multilingual environment of the European Union presenting multiple language ontologies and multiple language NSM.[3]

3 GETTING STARTED

3.1 How to Formulate an Ontology Query

The interface contains two sub-frames. One is for composing the 'Query', the other one is for sending the request. At this stage of the Multilingual Natural Language Query Interface, in the first query frame the user is free to compose requests involving ontology concepts, an NSM expression and a bounding box, as follows

1. Either one single ontology concept may be selected to give a simple list of resources that are semantically related to that concept, or two concepts may be selected as part of an NSM query (see point 2 below). The ontology concepts may be selected from either the **SBA categories and subcategories** in a list (see section 2.1.3 on 'selecting by SBA categories and subcategories') or the **GEMET** ontology in an appropriate client-interface

[1] http://www.eionet.europa.eu/gemet
[2] http://www.geoportal.org/web/guest/geo_home
[3] At present the NSM part of the interface has only been implemented for English. More languages will be added soon.

(see section 2.1.2 on 'selecting by GEMET concepts'), in a language of preference displayed in the box to the left of the screen.

2. The user may compose a natural language (NSM) expression to formulate a spatial relation between the two geographic features defined by the ontology concepts. For example

   - <GEMET:Forest Fires> BE(SOMEWHERE) INSIDE THIS <GEMET: drought>
   - PART OF THIS <GEMET:Forest ecosystem> BE(SOMEWHERE) INSIDE PART OF THIS <SBA:Migratory Species>

3. The user may select a map bounding box on the Google map displayed at the bottom of the interface, to retrieve only results within that area, that also comply to the selections described in points 1 and 2 above.

3.1.1 Selecting an Ontology Concept Using a Language of Preference

This procedure is compulsory if resources are to be discovered. It involves the selection of a concept from one of the two ontologies displayed on the screen - the GEOSS SBAs and GEMET.

The default ontology appearing automatically on screen when the page is loaded for the first time (at the centre of the green interface) is the SBAs, but GEMET can be chosen by clicking on the 'Ontology' drop down list to select another ontology, upon which a list of categories and (indented) subcategories to be selected appears.

The language used for the selected ontology may also be selected using the selection box at the left-hand side of the interface. Currently, GEMET supports up to 28 languages, while the SBAs support English, Italian, French, Spanish and Slovenian.

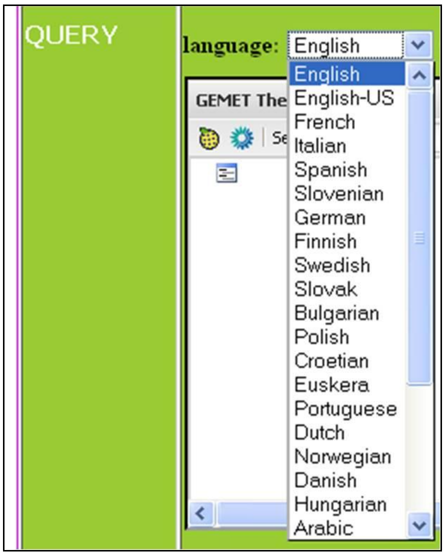

Figure 2: Language drop down list

Since the procedures for selecting concepts for GEMET and the SBA concepts are slightly different, they are described separately below.

3.1.2 *Selecting GEMET Concepts Using a Language of Preference*

If the user wishes to select a concept from GEMET (Figure 3 below), the user has two options. S/he could either select a language and type text in that language, say Italian, in the 'search' field thus clicking on the magnifying glass to find it (Figure 3) or alternatively s/he could click on one of the two left-hand side icons to see the concept lists provided by the ontology (Figure 4).

Figure 3: GEMET text box selection of concepts

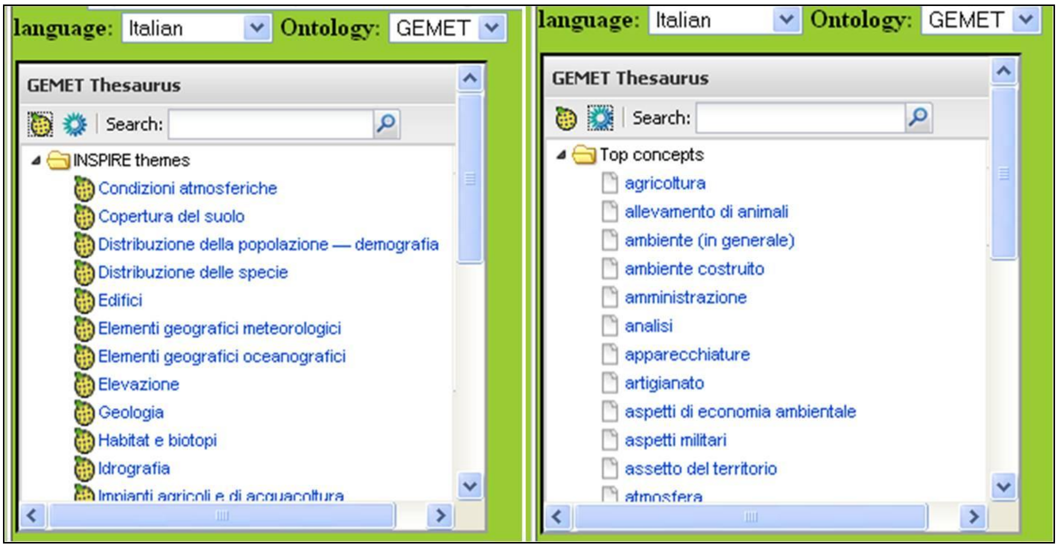

Figure 4: On the left the GEMET INSPIRE Themes, on the right the GEMET top concepts

The user can select concepts by browsing or by entering a search string in the text box. Browsing can be done through one of two top level hierarchies:

- The GEMET top concepts, which are displayed if the user clicks on the GEMET logo .

- The INSPIRE themes, which are displayed if the user clicks on the pineapple logo . The INSPIRE themes are a set of themes developed under the INSPIRE (Spatial Data Infrastructure in Europe) initiative. [4]

Figure 5: 'Use' of a concept in the GEMET client

Once the concept of interest, in the preferred language, has been selected the user can click on the button 'Use', as in Figure 5 above to add it to the current query.

Only when this last action is accomplished can the user click on 'send' in the 'Send Request' frame of the interface and wait for the resources response (see Figure 6). If the user clicks on 'send', s/he will receive back resources related to that concept and nothing else.

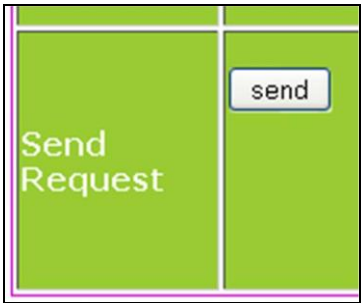

Figure 6: *Sending the request to the WPS*

3.1.3 Selecting by GEOSS SBAs in a Language of Preference

Should, instead, the user want to use a concept from the SBA categories and subcategories, this is much easier since s/he only has to browse it in a language of convenience, say French, in the list provided by the SBAs box (see Figure 7).

[4] http://inspire.jrc.ec.europa.eu/

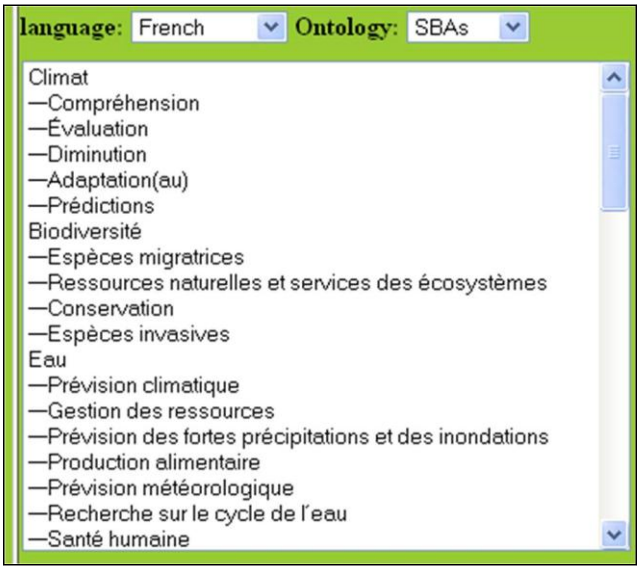

Figure 7: SBA categories and subcategories

As for the GEMET selection, once the concept has been selected by simply clicking on it, again the user has to then click on the 'send' button and wait a few seconds for a response.

Note that the response may take a few minutes to appear, depending on the number of semantically related resources that were found.

For further information please refer to the Help! Button 

above the send button for html instructions.

3.2 How to Formulate a Natural Language (NSM) Expression

3.2.1 What is NSM and Why Use It?

The simple selection of an ontology concept may be augmented with a natural language query to express a particular spatial relation. The interface uses a restricted natural language called Natural Semantic Metalanguage (NSM) for this purpose. NSM is a set of 63 linguistic primitives thought to be universal across a vast variety of languages (empirically demonstrated).[5] It is the result of a linguistic study spanning 40 years and initiated by the linguists Andrzej Bogusławski (Bogusławski, 1966) and Anna Wierzbicka during the '60s (Wierzbicka, 1980, 2002) and developed further by other linguists such as Cliff Goddard (Goddard, 2008).

We apply NSM to the geospatial field to formulate spatial expressions that match geospatial standard spatial operators (listed below) compliant to the Open Geospatial Consortium (OGC) standards but assisting the user to express them in different and non-technical ways. The supported spatial relations (and a simplified explanation of each) are exposed below:

[5] http://www.une.edu.au/bcss/linguistics/nsm/semantics-in-brief.php

OGC spatial operator	Simplified explanation
Within	Geometry X is contained in Geometry Y
Contains	Geometry Y contains Geometry X
Overlaps	The intersection of Geometry X and Geometry Y is a subset constituted by internal parts of X and internal parts of Y
Touches	The external boundary of Geometry X touches the external boundary of Geometry Y
Equals	Geometry X and Geometry Y occupy exactly the same place
Disjoint	No part of Geometry X and Geometry Y are in the same place

Figure 8: OGC Spatial Operators and their definitions

The idea behind the use of NSM in a geospatial context is to give the user the possibility to express the above mentioned spatial relations in a more intuitive and less technical way, and also to allow for the possibility to construct queries in different languages (still to be developed). Interested readers may refer to Stock and Cialone (Stock & Cialone, 2011) for more details on how NSM allows users to describe spatial relations in a way that allows linguistic and individual differences in interpretation to be expressed. To any of these spatial operators in fact, a set of possible NSM expressions has been grammatically and semantically mapped so that when the user formulates the spatial query between two geographic objectives, s/he will automatically trigger a standard spatial operator (of the ones above) that is understandable for the machine to read, parse and process geometrically.

3.2.2 How To Formulate an NSM Expression

To formulate an NSM expression is very easy! It does not require any linguistic or technical geospatial knowledge (that is the whole purpose in using NSM), but only intuition.

With NSM you can use 63 words/phrases as in Figure 9 below:

Substantives:	I/ME, YOU, SOMEONE, PEOPLE,SOMETHING/THING, BODY
Relational substantives:	KIND, PART
Determiners:	THIS, THE SAME/SAME, OTHER/ELSE
Quantifiers:	ONE, TWO, SOME, ALL, MUCH/MANY
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORDS, TRUE
Actions, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH
Location, existence, possession, specification:	BE(SOMEWHERE),THERE IS / EXIST, HAVE, BE(SOMEONE/ SOMETHING)
Life and death:	LIVE, DIE
Time:	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space:	WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
"Logical" concepts:	NOT, MAYBE, CAN, BECAUSE, IF
Intensifier:	VERY, MORE
Similarity:	LIKE/AS

Figure 9: Table of NSM primitives that can be used to formulate a spatial expression

The 63 primitives above can be used to formulate a natural language expression.

An NSM expression is a combination of primitives. Users can combine NSM primitives to formulate a spatial concept. Some concepts, due to their inner conceptual complexity such as *crosses* can mean many things for different speakers and in different languages and they need to be expressed in a more primitive way as follows:

- PART OF THIS <ROAD> IS INSIDE THIS <PARK>
  PART OF THIS SAME <ROAD> IS NOT INSIDE THIS SAME <PARK> ON THIS SIDE
  PART OF THIS SAME <ROAD> IS NOT INSIDE THIS SAME <PARK> ON THE OTHER SIDE

For this phase of implementation of the interface we will deal with relatively easy spatial concepts and easier explications that do not require long statements or painstaking exercises to be expressed, but only require a simple clause.

The words displayed in Figure 9 above do not include any grammatical form (e.g., conjugations, cases, genre, number). This means that the user must select the word that is the basis of the word of interest, once tense, case, number etc has been removed. For instance if one wants to say in English:

(1) **X** is where **Y** lives

or

(2) **X** does not live where **Y** lives

The resulting NSM explications should be respectively as follows:

(1) **X** BE(SOMEWHERE) WHERE/PLACE **Y** LIVE

(2) **X** NOT LIVE (or LIVE NOT) WHERE/PLACE **Y** LIVE

And so on.

What is more, some primes have two forms. These two forms are two linguistic variants of the same concept (THERE IS/EXIST) but that could be used in different contexts. For example it is possible in English to say

(1) **X** lives in this place

But not

(2) **X** lives in this where

And it is possible to say

(3) Where **X** lives

But not

(4) Place **X** lives (at least not without using where again as in 'place **where X** lives')

The concept of location is the same but the word used in common parlance is different. So, in NSM two single variants are presented, separated by a slash in one only prime so that the users could use them in different contexts without it being a mistake and meanwhile understanding they are expressing the same concept. At this stage, a limited set of spatial queries are supported by the demonstrator (see Appendix A for details of these). However, this set will be expanded in the coming months.

3.2.3 How To Formulate an NSM Expression Using the WPS MNL Interface

On the right-hand side of the screen there is a box (Figure 10 below) with a list of the natural language primitives in NSM (as in Figure 9 above), which magnify when the mouse moves over them. As already mentioned, these NSM primitives can be used to express a spatial query that can be understood and processed by the system. In this way, the user can express a spatial query between two ontology concepts to narrow down his or her search.

Figure 10: NSM Terms Box

If for instance, a thematic biodiversity user wants to know about resources concerning ‘forest fires within Europe’. S/He would first select the concept ‘forest fires’, then add some NSM primitives to compose the spatial query as shown in the previous section, and then select ‘forest’ from the two ontologies (as explained above). NSM primitives can be used to describe a range of different spatial relations in various ways. For example:

- FOREST FIRE INSIDE FOREST ECOSYSTEM.

This provides some flexibility for the user in expressing things in a way that suits him or herself.

The final spatial expression will appear in the middle box (NSM, Figure 11 below) and it is ready to be sent.

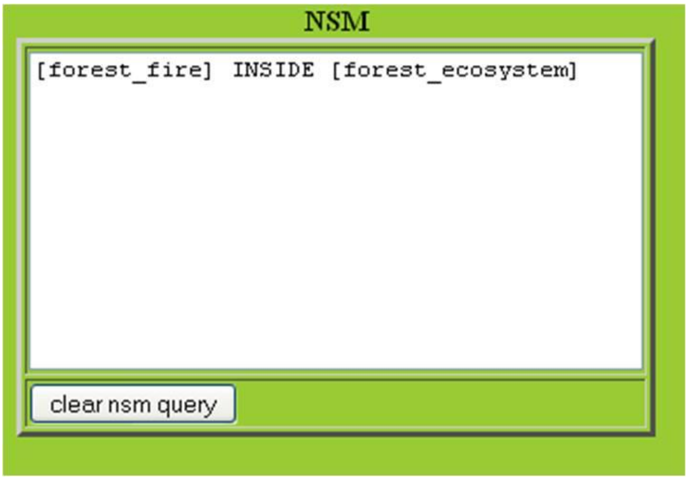

Figure 11: NSM spatial query with ontology concepts

3.3 Selecting a Map Bounding Box

The query can also be confined to a particular geographic region (e.g. in Spain), with the selection of a geometrical box using the Google map. The user may choose between a simple map, a satellite or a hybrid image (see Figure 12).

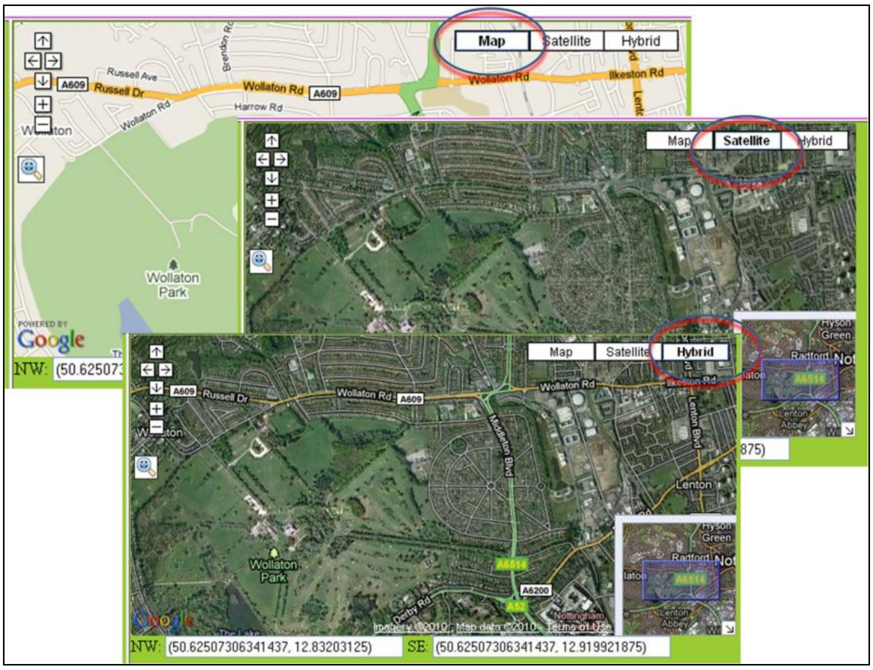

Figure 12: Screen-dumps of the three different types of map

The area of interest can be specified by simply clicking on the left-hand side magnifying glass icon onto the map and drawing a box by dragging the cursor to the preferred size of the geometrical area to identify it. The map can be zoomed in or out using the plus and minus buttons at the left-hand side of the map and the image can also be moved with the up/down and sideways arrows (see Figure 13). For the user who wants to know the numerical coordinates of the box can see them in the two text boxes below the map.

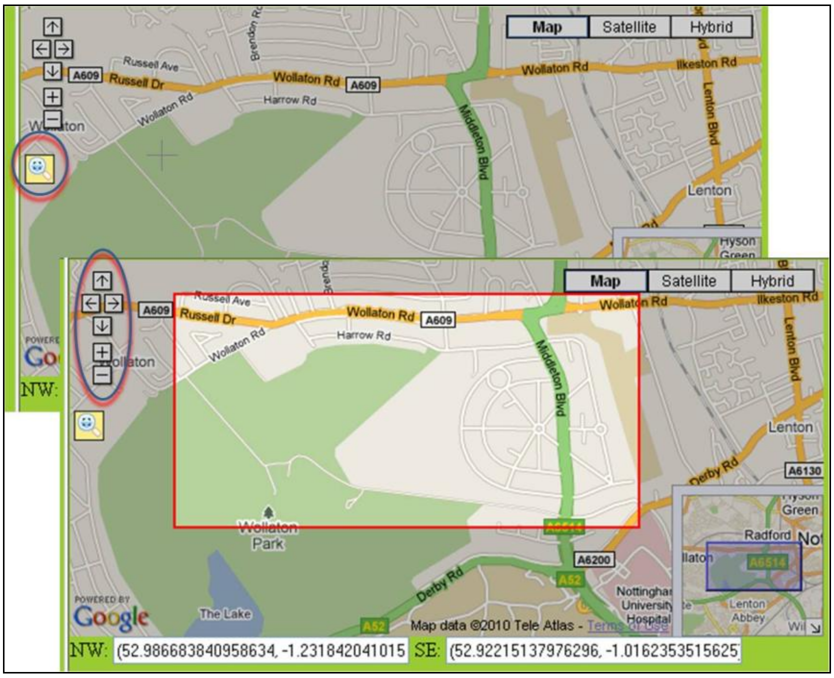

Figure 13: An example bounding box screen-shots of areas around Nottingham

There is also another (optional) box, at the very bottom right corner of the Google map, assisting the user to visualize a more general geographical context of orientation to which the area selected by the user belongs, (for instance the continent on its own as in Figure 14, depending on the scale of the user's magnification).

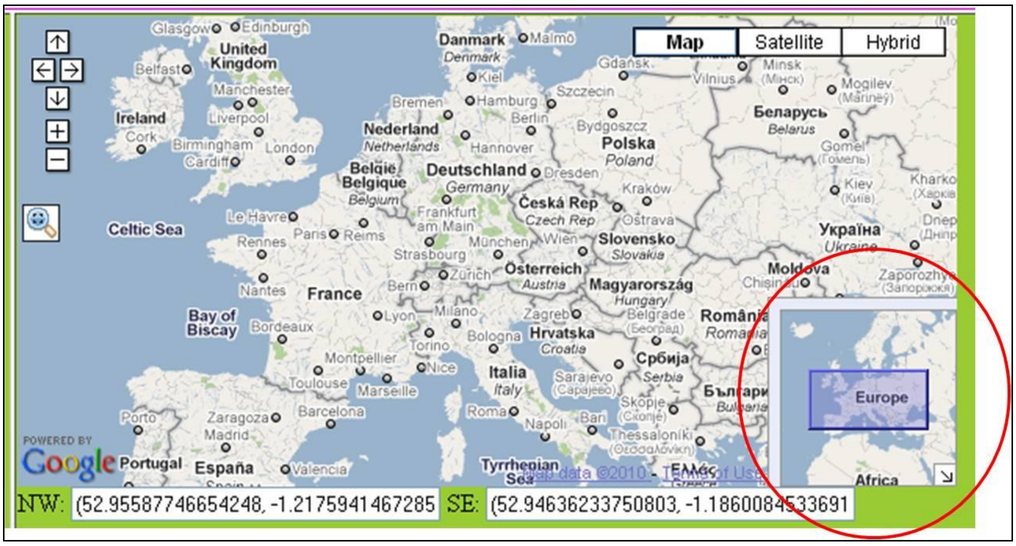

Figure 14: In the frame the orientation scale box is underlined

The choice of the coordinates can be easily reset through the provided 'reset Bbox' button Reset BBox at the side of the coordinate boxes, to start again with a new visual browsing.
The selection of a bounding box is optional, mainly allowing a particular geographic area to be searched. It is mandatory to select at least one ontology concept (described above).

4 RESULT

4.1 What Should the Response of the Natural Language Query Be?

4.1.1 One Ontology Concept Selected

If the request only contains **one** ontology concept (and optionally a set of coordinates), once the request has been sent, the system will generate a 'wait while working' message (Figure 16 below) and once the process terminates the result shows:

- The resources titles' links (e.g., cluster1)
- The language the Metadata has been written in.
- A brief abstract describing which type of resource it is (e.g. This GBIF OGC WMS layer provides over 7 million taxon occurrence records for the African continent including mammals, birds, reptiles, amphibia and insect [...]).
- The keywords attached to the resource.
- The semantic distance this resource is from the selection of the user (0,1,2,3,4. In order of semantic similarity), which tells (as well as the position of the resource in the list) the user how similar the resource is to his or her original selection.

The list of results will include not only resources that are tagged with the selected concept, but also resources that are tagged with other concepts that are semantically related to the selected concept. The resources (web services) have been previously tagged with ontology concepts, and those concepts have been mapped to each other so as to establish associative or hierarchical relations between different ontologies to expand, from a conceptual point of view, the simple query typed by the user. Therefore, the response corresponds to the combination of the geographic location and the ontology concept(s) selected (and semantically similar concepts). These are displayed under a numbered list from which the user can click on the most relevant according to his/her requirements (see Figure 15 below). Clicking on the resources metadata the user is led to the resource (web service) itself with the information required.[6]

Figure 15: Waiting Message

[6] At the moment we do not provide a service of visualization for the individual web services. This will follow.

Resource List

[GEMET: mammal]

1.
   Resource: cluster1
   Language: Unknown
   Abstract: This GBIF OGC WMS layer provides over 7 million taxon occurrence records for the African continent including mammals, birds, reptiles, amphibia and insecta. Occurrences are clustered into 1 degree cells.
   Keywords: mammal, bird, Reptile, Amphibian, Insect, species, Migratory species, Invasive species,
   Semantic Distance: 0
2.
   Resource: cluster10
   Language: Unknown
   Abstract: This GBIF OGC WMS layer provides over 7 million taxon occurrence records for the African continent including mammals, birds, reptiles, amphibia and insecta. Occurrences are clustered into 0.1 degree cells.
   Keywords: mammal, bird, Reptile, Amphibian, Insect, species, Migratory species, Invasive species,
   Semantic Distance: 0
3.
   Resource: cluster100
   Language: Unknown
   Abstract: This GBIF OGC WMS layer provides over 7 million taxon occurrence records for the African continent including mammals, birds, reptiles, amphibia and insecta. Occurrences are clustered into 0.01 degree cells.
   Keywords: mammal, bird, Reptile, Amphibian, Insect, species, Migratory species, Invasive species,
   Semantic Distance: 0
4.
   Resource: occurrence_full
   Language: Unknown
   Abstract: This GBIF OGC WMS layer provides over 7 million taxon occurrence records for the African continent including mammals, birds, reptiles, amphibia and insecta.
   Keywords: mammal, bird, Reptile, Amphibian, Insect, species, Migratory species, Invasive species,
   Semantic Distance: 0

Figure 16: MNL WPS resource discovery response single

4.1.2 Two Ontology Concepts & NSM Expression Selected

If the request only contains **two** ontology concepts and an NSM expression (and optionally a set of coordinates), once the request has been sent, the system will generate a 'wait while working' message (Figure 15 above) and once the process completes, the result shows the OGC spatial operator(s) (see Figure 17) that the user query was mapped to, together with pairs of resources that are tagged with the selected ontology concepts (or semantically similar concepts), and that have the spatial relation to one another expressed by the NSM query. This means that if for example the first query with one only ontology concept included <drought> (inside an optionally defined bounding box), the system will only retrieve all those resources related to drought (in that given bounding box); if instead the user is interested in retrieving all those resources related to <species> living inside <drought>, the system will retrieve all the biodiversity theme related resources whose features are included in a geographical area that is located inside another geographical area tagged as a drought affected area.

The response message for couples of resources (A, B) whose features are included in bounded boxes that have been filtered through the OGC spatial operators can be seen below (this is a general example for <species> inside <land cover>). The result is still ranked in order of importance.

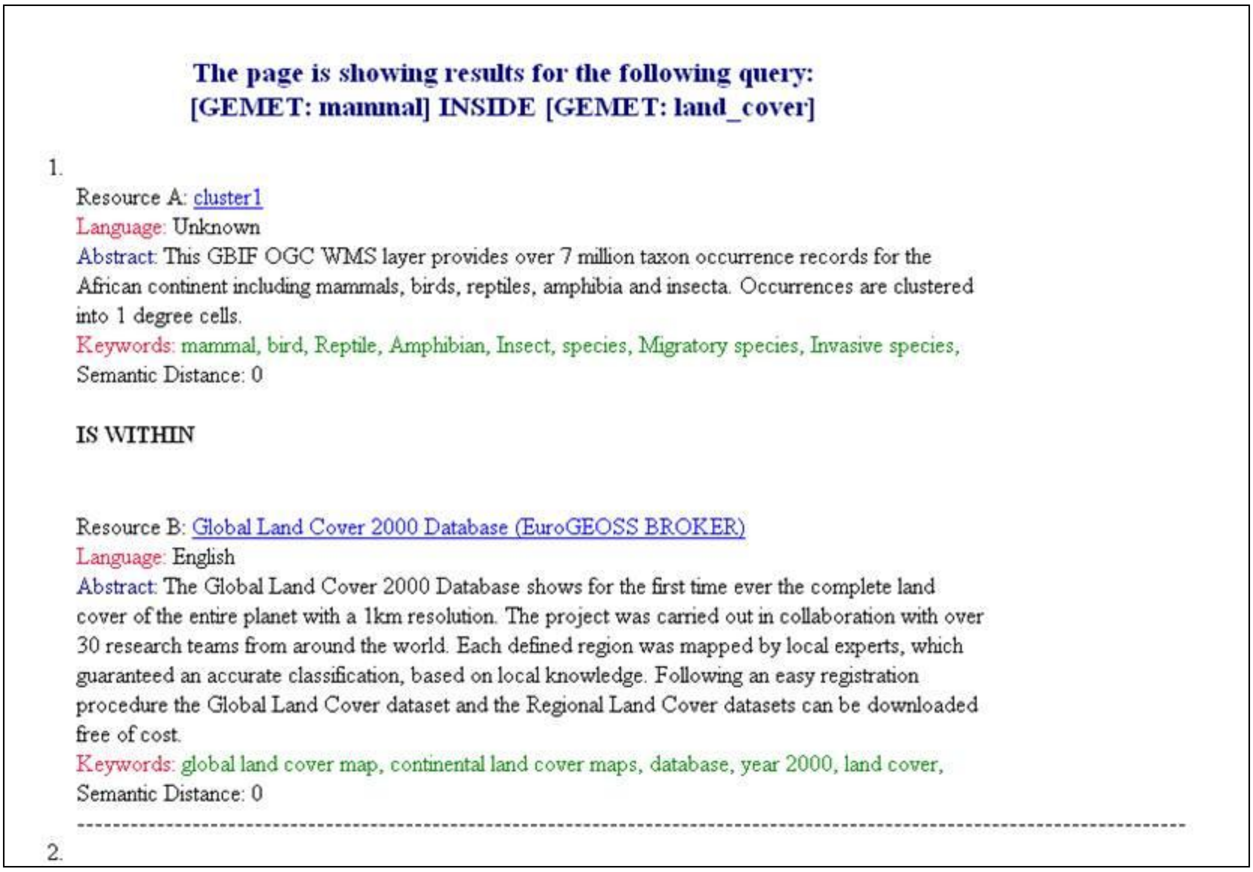

Figure 17: MNL WPS resource discovery response with spatial operator

4.1.3 No-Resources Found !

If no resources are found the interface will show another message as in Figure 18 below and the possibility to go back to the interface page (circled in red, below).

Figure 18: No results message

5 SUMMARY/CONCLUSIONS

In this practical set of instructions we have shown

- how to select an ontology concept;
- how to optionally formulate and specify a spatial relation (in the form of an NSM expression) between two ontology concepts by using the interface and
- how to optionally specify a bounding box by drawing coordinates on a map.

The purpose of the interface is to formulate a multi-lingual and meaningful geospatial query that will be able to retrieve resources (mainly web services) related to Biodiversity, Drought and Forestry compliant with the scope of the EuroGEOSS project (2009-2012). The goal of the work is to enable queries to be performed in an intuitive way, preserving individual and language-specific modes of expression, without requiring expert knowledge or understanding of spatial relations and their mathematical meanings. This demonstrator takes a first step towards the goal, supporting a range of different user queries (see Appendix A), and will be expanded in the future to support a growing range of queries and methods of expression.

REFERENCES

Bogusławski, A.: On Semantic primitives and meaningfulness.In Signs Language and Culture.Proceedings of a Conference Held in Kazimierz, The Hague: Mouton (1966).

Cialone C., Stock K., 'Domain Ontologies to Support EuroGEOSS', internal report (final).

Cialone C., Stock K., 'Multilingual Knowledge Systems. The EuroGEOSS' Case Study:GEOSS Societal Benefit Areas Translations for Italian, Spanish and French', (forthcoming 2011).

Goddard C., *Cross-Linguistic Semantics*. John Benjamins Publishing Company. Amsterdam/Philadelphia, 2008.

Gomez-Perez A., Corcho-Garcia O., Fernandez-Lopez M..*Ontological Engineering: With Examples from the areas of knowledge Management, e-commerce and the Semantic Web*, Springer-Verlag London Limited, 2004.

Euzenat J., Shvaiko P.. *Ontology matching*, Springer-Verlag Berlin Heidelberg, 2007.

Nebert D., Whiteside A., Panagiotis (Peter) Vretanos. 'OpenGIS® Catalogue Services Specification', OGC 07-006r1, 2007, version 2.0.2

Panagiotis A. Vretanos, 'Open GIS® Filter Encoding Implementation Specification', OGC 04-095, 2005.

Peeters B., *Semantic Primes and Universal Grammar : Empirical Evidence from the Romance Languages.* John Benjamins Publishing Company 2006.

Schut P. 'OpenGIS® Web Processing Service', OGC 05-007r7, 2007, version 1.0.0

Stock K., Cialone C., 'An Approach to the Management of Multiple Aligned Ontologies for a Geospatial Earth Observation System', GeoS2011 (2011).

Stock,K., Cialone, C.: Universality, Language-Variability and Individuality: Defining Linguistic Building Blocks for Spatial Relations, COSIT (forthcoming 2011)

Talmy L., *Toward a Cognitive Semantics*, volumes I and II, MIT Press, 2000, http://linguistics.buffalo.edu/people/faculty/talmy/talmy.html

Wierzbicka, A.: Lingua Mentalis: The Semantics of Natural Language. Sydney, New York, London, Toronto, San Francisco (1980).

Wierzbicka A., Goddard C., *Meaning and Universal Grammar: Theory and Empirical Foundings*, vol. I,II. John Benjamins Publishing Compay, Amsterdam/Philadelphia, 2002.

Goddard, C.: Cross Linguistic Semantics, Amsterdam/Philadelphia: John Benjamin Publishing Company (2008)

APPENDIX A: CURRENTLY SUPPORTED QUERIES

The demonstrator currently supports a limited range of queries, both in regard to the ontology concepts that return results and the spatial queries that are supported. This Appendix describes the queries that are currently supported.

A.1 Ontology Concepts

Only a limited number of resources have been tagged with ontology concepts as part of the EuroGEOSS project. Thus the user must select these concepts in their query in order to get results. The concepts that have been used to tag resources to date are:

- GEMET:
  - atmospheric emission;
  - greenhouse gas
  - air pollutant
  - forest fire
  - land cover
  - ecosystem
  - forest ecosystem
  - mammal
  - bird
  - reptile
  - amphibian
  - species
  - migratory species
  - biodiversity
  - water (substance)
  - energy
  - agriculture
  - drought
- SBAs:
  - ecosystems
  - migratory species
  - invasive species
  - energy
  - agriculture
  - water
  - biodiversity
  - land, river, coast and ocean management
  - drought prediction.

A.2 Spatial Queries

For the demonstrator, we are temporary supporting a limited range of NSM explications. These are the ones that describe spatial relations in the ways that appeared most commonly in our user experiments. It is our intention, though, to add capability for a wider range of spatial query descriptions in the near future. In these queries, X and Y are used in place of the ontology concepts.

In the queries below, both the version that includes the case, plurality and other grammatical variations and the version in which the grammatical base for each primitive is selected from the drop down list in the user interface (see Section 2.2.2) are shown to assist users in understanding the meaning.

In most cases the variations on these queries can be mixed and matched (for example, NOW or HERE can be added to the end of any query without the results being affected, if this is what the user chooses to write).

Grammatically Specific Version	Drop-down box selection version
X IS INSIDE Y	X BE(SOMEWHERE) INSIDE Y X THERE IS/EXIST INSIDE Y
X IS INSIDE Y NOW	X BE(SOMEWHERE) INSIDE Y NOW
X LIVES INSIDE Y	X LIVES INSIDE Y
X LIVES INSIDE Y IN THIS PLACE	X LIVES INSIDE Y THIS WHERE/PLACE
X HAS Y INSIDE	X HAVE/BELONG Y INSIDE
X WAS INSIDE Y BEFORE	X BE(SOMEWHERE) INSIDE Y BEFORE
X WAS INSIDE Y BEFORE NOW	X BE(SOMEWHERE) INSIDE Y BEFORE NOW
THIS X IS INSIDE THIS Y	THIS X BE(SOMEWHERE) INSIDE THIS Y
X TOUCHES Y[7]	X TOUCHES Y[8]
PART OF X IS INSIDE PART OF Y	PART OF X THERE IS/EXIST INSIDE PART OF Y
PART OF X IS INSIDE PART OF Y HERE	PART OF X THERE IS/EXIST INSIDE PART OF Y HERE
ALL PARTS OF X ARE NOT IN THE SAME PLACE AS ALL PARTS OF Y	ALL PART OF X BE(SOMEWHERE) NOT THE SAME PLACE ALL PART OF Y
ALL PARTS OF X ARE IN THE SAME PLACE AS ALL PARTS OF Y	ALL PART OF X BE(SOMEWHERE) THE SAME PLACE ALL PART OF Y
X IS IN THE SAME PLACE AS Y	X THERE IS/EXIST SAME WHERE/PLACE Y

[7] But there are currently no tagged pairs of resources with bounding boxes for which ‘touches’ is true.
[8] But there are currently no tagged pairs of resources with bounding boxes for which ‘touches’ is true.