



Data, Information and Process Integration  
with Semantic Web Services

## **DIP**

*Data, Information and Process Integration with Semantic Web Services*

**FP6 - 507483**

Deliverable

# **WP 9: Case Study eGovernment D9.6 GIS Requirements Specification**

Leticia Gutiérrez (ECC)

Mary Rowlatt (ECC)

Rob Davies (ECC)

Alexander Wahler (NIWA)

Bernhard Peissl (NIWA)

May 18<sup>th</sup>, 2005



## SUMMARY

This deliverable describes the requirements of Essex County Council (ECC) in relation to a GIS prototype enhanced by SWS technology.

First, information about how Spatially Related Data (SRD) is managed in ECC has been collected by means of interviews with the relevant people from different departments in ECC. The DIP project and the potential benefits of SWS in their work were explained to interviewees, after which the necessary information about their work and requirements solicited.

Most of the interviewees referred to the **ViewEssex** project; a GIS system which ECC is currently creating to aggregate much of the SRD that is distributed across the organisation. This repository will be examined by the Essex DIP team to assess how much of the data held in it is relevant to the DIP prototype.

Further interviews will be held within the next two months in order to establish a heterogeneous range of data sources. Efforts will be made to interview representatives of external organizations relevant to the use case scenario (see below), for example police, fire-workers, airport managers, flight companies, etc) and to encourage them to participate in the project.

Based on the interview results a use case scenario has been chosen as the focus of the work, namely the area around London Stansted airport.

Following this, the prototype architecture was designed. It is divided into several layers. This has been done with the help of ECC GIS experts and the ideas of other technical partners in WP9. A work plan has also been produced, which details the tasks to be done, their deadlines and who is responsible for their accomplishment.

Although it is out of the scope of a requirements specification document; we have also attempted to weigh up the case for using IRS-III technology or WSMX tools to implement the prototype. It has been decided provisionally to use IRS-III in a first version of the prototype, and to switch to WSMX when this is ready in a second version of the prototype. The consequences of this change (for example, translation of the ontology from OCML into WSML) have been taken into account in creating the work plan and the deadlines.

A broad audience is targeted by this deliverable.

- Technical partners should read it, because it is the basis for future discussion about the prototype which will be implemented.
- Specialists within ECC should read this document to understand the objectives of the use cases and to obtain basic knowledge about DIP, why ECC is involved in it and how DIP technology will be tested in a ‘real world’ application.
- It may also be useful for other use-case partners (WP8, WP10) to read this document, in order to compare their scenarios and prototype architectures with WP9. This could be of great benefit for all the use-case partners in sharing ideas and mutually suggesting ways of improving systems.
- Finally, Technical WP partners should read this document to improve understanding of the role of individual WSMO features [23] in a ‘real world’ implementation (e.g.: discovery, mediation, composition) and whether they can be provided in time.

Finally, it is noted that this deliverable contributes to one of DIP’s main goals: the WS architecture. This document attempts to set out or illustrate: how the SWS technology can be used in a real world system, show SWS components fit into a system, how to get access to data sources, how to create WS on top of them, how to add semantics to plain WS, and how the user can discover and invoke SWS by invoking some goals through a friendly end-user application.

## Document Information

<b>IST Number</b>	<b>Project</b> FP6 – 507483	<b>Acronym</b>	DIP
<b>Full title</b>	Data, Information, and Process Integration with Semantic Web Services		
<b>Project URL</b>	<a href="http://dip.semanticweb.org">http://dip.semanticweb.org</a>		
<b>Document URL</b>			
<b>EU Project officer</b>	Kai Tullius		

<b>Deliverable</b>	<b>Number</b> 9.6	<b>Title</b>	GIS requirements specification
<b>Work package</b>	<b>Number</b> 9	<b>Title</b>	Case study eGovernment

<b>Date of delivery</b>	<b>Contractual</b>	M 15	<b>Actual</b>	26-May-05
<b>Status</b>	Version. 0.8		final <input checked="" type="checkbox"/>	
<b>Nature</b>	Prototype <input type="checkbox"/> Report <input checked="" type="checkbox"/> Dissemination <input type="checkbox"/>			
<b>Dissemination Level</b>	Public <input type="checkbox"/> Consortium <input checked="" type="checkbox"/>			


<b>Authors (Partner)</b>	Alex Wahler (NIWA), Bernhard Peissl (NIWA), John Domingue (OU), Leticia Gutierrez (ECC), Liliana Cabral (OU), Mary Rowlatt (ECC), Rob Davies (ECC)			
<b>Responsible Author</b>	Leticia Gutierrez		<b>Email</b>	<a href="mailto:Leticia.gutierrez@essexcc.gov.uk">Leticia.gutierrez@essexcc.gov.uk</a>
	<b>Partner</b>	ECC	<b>Phone</b>	01245 436185



<b>Abstract (for dissemination)</b>	This document illustrates the requirements specifications of the GIS prototype WP9 will create in the forthcoming months.
<b>Keywords</b>	Prototype, Ontology, Geographic Information Systems (GIS), Ontology-Driven Geographic Information Services (ODGIS) , Spatially-related data (SRD), raster, vector, Semantic Web Services (SWS), Semantic Web (SW), e-Government

<b>Version Log</b>			
<b>Issue Date</b>	<b>Rev No.</b>	<b>Author</b>	<b>Change</b>
02.02.2005	001	Leticia Gutiérrez (ECC) RobDavies (ECC) Mary Rowlatt (ECC),	Initial version of the Document
09.02.2005	002	Leticia Gutierrez (ECC)	Add functional and non-functional requirements
21.02.2005	003	Leticia Gutierrez (ECC) Alexander Wahler (NIWA) Bernhard Peissl (NIWA) Jon Bryant (BT Consulting & Systems Integration- ECC)	Add Alexander Wahler and Jon Bryant's suggestions about GIS notions to the deliverable
21.03.2005	004	Leticia Gutierrez (ECC)	Use case and sequence diagrams
24.03.2005		Rob Davies (ECC) Mary Rowlatt (ECC) Leticia Gutierrez (ECC)	Scenario and interviews description
29.03.2005	005	Leticia Gutierrez (ECC)	Functional requirements review
31.03.2005	006	Rob Davies (ECC) Mary Rowlatt (ECC) Leticia Gutierrez (ECC)	Objectives definition
14.04.2005	007	Alexander Wahler (NIWA) Bernhard Peissl (NIWA)	Informal review and comments
04.05.2005	008	Leticia Gutierrez (ECC) Sigurd Harald (DERI) Rob Davies (ECC) Silvestre Losada (ISOCO) Alistair Duke (BT)	Deliverable reviewed by Sigurd Harald, Rob Davies, Silvestre Losada and Alistair Duke All comments added to the deliverable

## Project Consortium Information

Partner	Acronym	Contact
National University of Ireland Galway	NUIG 	Prof. Dr. Christoph Bussler Digital Enterprise Research Institute (DERI) National University of Ireland, Galway Galway Ireland Email: <a href="mailto:chris.bussler@deri.org">chris.bussler@deri.org</a> Tel: +353 91 512460
Fundacion De La Innovacion.Bankinter	Bankinter 	Monica Martinez Montes Fundacion de la Innovacion. BankInter Paseo Castellana, 29 28046 Madrid, Spain Email: <a href="mailto:mmtnez@bankinter.es">mmtnez@bankinter.es</a> Tel: 916234238
Berlecon Research GmbH	Berlecon 	Dr. Thorsten Wichmann Berlecon Research GmbH Oranienburger Str. 32 10117 Berlin, Germany Email: <a href="mailto:tw@berlecon.de">tw@berlecon.de</a> Tel: +49 30 2852960
British Telecommunications Plc.	BT 	Dr John Davies BT Exact (Orion Floor 5 pp12) Adastral Park Martlesham Ipswich IP5 3RE, United Kingdom Email: <a href="mailto:john.nj.davies@bt.com">john.nj.davies@bt.com</a> Tel: +44 1473 609583
Swiss Federal Institute of Technology, Lausanne	EPFL 	Prof. Karl Aberer Distributed Information Systems Laboratory École Polytechnique Fédérale de Lausanne Bât. PSE-A 1015 Lausanne, Switzerland Email : <a href="mailto:Karl.Aberer@epfl.ch">Karl.Aberer@epfl.ch</a> Tel: +41 21 693 4679
Essex County Council	Essex 	Mary Rowlett, Essex County Council PO Box 11, County Hall, Duke Street Chelmsford, Essex, CM1 1LX United Kingdom. Email: <a href="mailto:maryr@essexcc.gov.uk">maryr@essexcc.gov.uk</a> Tel: +44 (0)1245 436524
Forschungszentrum Informatik	FZI 	Andreas Abecker Forschungszentrum Informatik Haid-und-Neu Strasse 10-14 76131 Karlsruhe Germany Email: <a href="mailto:abecker@fzi.de">abecker@fzi.de</a> Tel: +49 721 9654 0

Partner	Acronym	Contact
Institut für Informatik, Leopold-Franzens Universität Innsbruck	UIBK 	Prof. Dieter Fensel Institute of computer science University of Innsbruck Technikerstr. 25 A-6020 Innsbruck, Austria Email: <a href="mailto:dieter.fensel@deri.org">dieter.fensel@deri.org</a> Tel: +43 512 5076485
ILOG SA	ILOG  Changing the rules of business	Christian de Sainte Marie 9 Rue de Verdun, 94253 Gentilly, France Email: <a href="mailto:csm@ilog.fr">csm@ilog.fr</a> Tel: +33 1 49082981
inubit AG	Inubit  the integration experts	Torsten Schmale inubit AG Lützowstraße 105-106 D-10785 Berlin Germany Email: <a href="mailto:ts@inubit.com">ts@inubit.com</a> Tel: +49 30726112 0
Intelligent Software Components, S.A.	iSOCO 	Dr. V. Richard Benjamins, Director R&D Intelligent Software Components, S.A. Pedro de Valdivia 10 28006 Madrid, Spain Email: <a href="mailto:rbenjamins@isoco.com">rbenjamins@isoco.com</a> Tel. +34 913 349 797
The Open University	OU 	Dr. John Domingue Knowledge Media Institute The Open University, Walton Hall Milton Keynes, MK7 6AA United Kingdom Email: <a href="mailto:j.b.domingue@open.ac.uk">j.b.domingue@open.ac.uk</a> Tel.: +44 1908 655014
SAP AG	SAP 	Dr. Elmar Dörner SAP Research, CEC Karlsruhe SAP AG Vincenz-Priessnitz-Str. 1 76131 Karlsruhe, Germany Email: <a href="mailto:elmar.dorner@sap.com">elmar.dorner@sap.com</a> Tel: +49 721 6902 31
Sirma AI Ltd.	Sirma  Ontotext Knowledge and Language Engineering Lab of Sirma	Atanas Kiryakov, Ontotext Lab, - Sirma AI EAD Office Express IT Centre, 3rd Floor 135 Tzarigradsko Chausse Sofia 1784, Bulgaria Email: <a href="mailto:atanas.kiryakov@sirma.bg">atanas.kiryakov@sirma.bg</a> Tel.: +359 2 9768 303

Partner	Acronym	Contact
Unicorn Solution Ltd.	Unicorn 	Jeff Eisenberg Unicorn Solutions Ltd, Malcha Technology Park 1 Jerusalem 96951 Israel Email: <a href="mailto:Jeff.Eisenberg@unicorn.com">Jeff.Eisenberg@unicorn.com</a> Tel.: +972 2 6491111
Vrije Universiteit Brussel	VUB 	Carlo Wouters Starlab- VUB Vrije Universiteit Brussel Pleinlaan 2, G-10 1050 Brussel ,Belgium Email: <a href="mailto:carlo.wouters@vub.ac.be">carlo.wouters@vub.ac.be</a> Tel.: +32 (0) 2 629 3719

---

## LIST OF KEY WORDS/ABBREVIATIONS

The following list catalogues the special terms that apply to the domain on this GIS-project, and other acronyms that should be clarified.

<b>BT-C&amp;SI</b>	(British Telecom Consulting & Systems Integration) IT company in partnership with the ECC
<b>DIP</b>	Data Information and Process integration within Semantic Web Services
<b>ECC</b>	Essex County Council
<b>GIS</b>	Geographical Information System. Computer software, hardware, data and people that store, analyse, manipulate and present information linked to a spatial location. [1]
<b>NIWA</b>	SME specialised in IT Services, located in Vienna
<b>ODGIS</b>	Ontology-Driven Geographic Information Services
<b>OU</b>	Open University
<b>SOAP</b>	Simple Object Access Protocol
<b>SRD</b>	Spatially Related Data
<b>SW</b>	Semantic Web
<b>SWS</b>	Semantic Web Service(s)
<b>UDDI</b>	Universal Description Discovery and Integration
<b>WP9</b>	Work Package 9 within DIP project: eGovernment case study
<b>WS</b>	Web service
<b>WSML</b>	Web Service Modelling Language
<b>WSMO</b>	Web Service Modelling Ontology
<b>WSMX</b>	Web Service Execution Environment

---

## TABLE OF CONTENTS

<b>SUMMARY</b> .....	<b>II</b>
<b>LIST OF KEY WORDS/ABBREVIATIONS</b> .....	<b>IX</b>
<b>TABLE OF CONTENTS</b> .....	<b>X</b>
<b>1 INTRODUCTION</b> .....	<b>13</b>
<b>2 PROBLEM DESCRIPTION</b> .....	<b>14</b>
2.1 ECC spatial data current situation .....	14
2.1.1 Current solutions .....	14
2.2 How to improve GIS using semantics .....	15
2.3 Building WS and SWS over a GIS .....	17
2.3.1 WS technology .....	17
2.3.2 Benefits of WS over a common repository .....	18
2.3.3 Benefits of SWS over WS .....	19
2.4 GIS Scenarios .....	21
2.5 Identifying key experts .....	21
2.6 Interviews .....	22
2.6.1 Description of interviews .....	22
2.6.2 Overview of interview results .....	22
2.7 Scenario selection .....	26
2.8 Potential systems to be merged .....	30
<b>3 PROJECT CONSTRAINTS</b> .....	<b>31</b>
3.1 Implementation Environment of the Current System.....	31
3.2 Ontology Constraints .....	31
3.3 WSMO operations required.....	32
<b>4 REQUIREMENTS</b> .....	<b>33</b>
4.1 Functional Requirements .....	33
4.1.1 Objectives .....	33
4.1.2 GIS systems integration.....	34
4.1.3 WS Requirements .....	34
4.1.4 Special WS Requirements .....	35
4.1.5 Graphical User Interface Requirements .....	35
4.1.6 Use case diagrams .....	40
4.2 Non-Functional Requirements.....	45
4.2.1 Accuracy.....	46

---

4.2.2 Integration into existing infrastructure .....	46
4.2.3 Scalability and Performance.....	47
4.2.4 Security and performance.....	47
4.2.5 Different access rights for service consumers .....	47
4.2.6 Availability .....	48
<b>5 PROTOTYPE .....</b>	<b>48</b>
5.1 Architecture .....	49
<b>6 CONCLUSION .....</b>	<b>50</b>
<b>REFERENCES .....</b>	<b>50</b>
<b>REFERENCES .....</b>	<b>51</b>
Annex 1 - State-of-the-art in Geographic Services .....	52
Open Geospatial Consortium .....	52
Other Current Core Technologies .....	53
Annex 2 - Questions' template for the interviews.....	55
▪ <i>General Questions (Real-world problems)</i> .....	55
▪ <i>More Technical Questions</i> .....	57
▪ <i>Very Technical Questions</i> .....	57
Annex 3 - Metadata schema.....	58

---

**LIST OF FIGURES**

Figure 1: Stansted map. Source: <a href="http://www.multimap.com">www.multimap.com</a> .....	27
Figure 2: Stansted map. Source: <i>ViewEssex</i> .....	28
Figure 3: Essex map. Source: <i>ViewEssex</i> .....	29
Figure 4: How DIP project fits into <i>ViewEssex</i> and mediates with other projects .....	31
Figure 5: Use case diagram .....	41
Figure 6: Sequence diagram of the ‘airport expansion consequences’ use case .....	43
Figure 7: Sequence diagram of the ‘housing expansion consequences’ use case – A ...	44
Figure 8: Sequence diagram of the ‘housing expansion consequences’ use case - B ....	44
Figure 9: Sequence diagram of the ‘simulate independent issues’ use case .....	45
Figure 10: System prototype .....	48
Figure 11: OpenGIS Abstract Specification Bookshelf, Topic 0 .....	53
Figure 12: Ontology schema example .....	58

**LIST OF TABLES**

Table 1: Summary table of interviews results .....	26
Table 2: Stansted expansion use case description table .....	42
Table 3: Housing expansion use case description table .....	43
Table 4: Expansion-independent use case description table .....	45

## 1 INTRODUCTION

This project will demonstrate how the Semantic Web (SW) and more concretely, Semantic Web Service (SWS) based systems; can be applied to improve spatially-related systems. The final product of this project will be a prototype developed in ECC with the support of NIWA, OU and BT.

The aim of this project is to demonstrate how SWS technology can be applied to a Geographic Information environment. ECC holds a large amount of Spatial-related Data (SRD) and manages a number of Geographical Information Systems (GIS). This wide distribution across the organization is the cause of a great deal of duplication of data and lack of communication among SRD holders. Through implementation of a SWS-GIS system all the data would be made available through the web, and it automated management would be made possible (discovery, composition and invocation).

This document provides the requirements specification, which is the part of the project during which the needs of the client are gathered and translated into a design of what the system must do. In this particular case, ECC is simultaneously the client and the developer. ECC personnel involved in the DIP project will act as the developer (together with OU, BT and NIWA). People interviewed so far belong to different departments in ECC, but in the near future more interviews will take place which may include people from external organizations (police, fire-workers, airport managers, flight companies, etc). The interviewees chosen are ECC personnel responsible for Geographical Information Systems (GIS) and other relevant data sources.

We would like to thank all the people that have kindly collaborated in this project, for their time and expertise sharing with us.

This document is divided as follows: Section 2 (Problem Description) describes the problem being approached: describes the state-of-the-art in Geographic Services, how ECC manages data currently, how SWS would improve data management within Essex, potential scenario applications and how these were identified within the organization. Section 3 (Project Constraints) provides information about the project constraints (about the implementation environment and ontology). Section 4 (Requirements) describes the requisites of the system (functional and non-functional), and depicts use case diagrams and sequence diagrams to clarify the information flow. Section 5 (Prototype) describes the architecture of the prototype, and finally, section 6 (Conclusion) summarises the content of the deliverable.

## 2 PROBLEM DESCRIPTION

An introduction to current Geographical Information Systems (GIS) standards and core technologies can be found in Annex 1.

### 2.1 ECC spatial data current situation

GIS are used to support decision making across a very wide range of functional areas within ECC. Typically, these are isolated systems, containing different kinds of heterogeneous data, for example: Highways and Transportation, Planning, Learning and Social Care, etc.

This heterogeneity has its origin in the way different organisations manage data: every organisation has its own view of the world and only stores the data that is relevant to achieve its goals, arranged in a way which is apparently easily manageable.

The lack of a common agreement about the structure of geographic information leads to various problems, such as different coordination standards, different mappings and limited knowledge of what information is available to ECC.

In addition, ECC still has a large number of processes which are manual or paper-based, or information systems which could use geospatial data but do not, for example, using address data as alpha-numeric fields but not geocoding them [1]

#### 2.1.1 Current solutions

ECC is working on a new project from the beginning of 2004 called **ViewEssex**: “A business case to develop an ECC and County-wide approach to managing spatial-data with quantified costs/benefits” [1]

**ViewEssex** is a big and ambitious project, which will deliver a central spatial data repository for the whole council and potentially partners such as district councils, and ultimately the general public (via the Internet). However there remains huge potential for greater exploitation of these information assets by understanding their semantic relationships.

**Example of semantics:**

*Using semantics, applications can “understand” what a road is: a place where vehicles pass through to go from one place to another. Different kinds of roads can be stated: highways, regional roads, national roads, etc. every one having different attributes (speed limit, pavement characteristics, access restrictions for certain types of vehicle. pedestrian access etc). Another characteristic of roads is their speed limit: with semantics, the speed limit can be stated for every road, as can the level of fines for cars if this is disregarded and the date by which payment must be made. The maximum number of vehicles per hour using the road in order to enable acceptable traffic flow can also be established.*

As ECC is leading the **ViewEssex** project and is also involved in WP9 of the DIP project (Data, Information and Process Integration with Semantic Web Services), this case study attempts to bring both together to develop a semantically-improved GIS system.

The objective of the proposed system is to create a SWS-enhanced GIS system which will exploit interoperability between spatial data spread out across ECC and its partners, with a minimum effort, supporting ECC, decision makers and eventually public end users.

## 2.2 How to improve GIS using semantics

Any information system can gain advantage from the use of semantics. The use of semantics in GIS, although not very well established, is being investigated in a number of research studies ([18][19][20][21][22])

The enhancement of GIS with Ontologies is sometimes known as ODGIS (Ontology-Driven GIS). An Ontology, which is an “*explicit and formalized specification of a conceptualisation*” [4], is created to define the semantics of the GIS, and is used afterwards as the backbone of the semantic-based system. This is explained in detail in later chapters of the document)

An Ontology represents an abstract and simplified view (vocabulary, relationships and logical rules) of the piece of reality it intends to represent. Ontologies (a relatively new technology which appeared in the 1990s) define an agreed common terminology by providing concepts, and relationships between this set of concepts. They are the means of **defining the concepts of the data exchanged without human intervention**. The main purpose of Ontologies within the SW is to provide meaning, enabling knowledge sharing and reuse between users and intelligent software agents.

Annex 3 depicts and Ontology schema relevant to the previously described example of roads. An Ontology depicts metadata: data about data. In our contexts it refers to machine understandable information for the web. (<http://www.w3.org/Metadatal/>). However this example is not a complete meta-model, as it only attempts to show how GIS elements and their intrinsic relations can be illustrated within an Ontology.

Some of the advantages of using ODGIS over plain GIS are:

- Enabling the performance of queries based on semantic values instead of syntax (*e.g: show all fields suitable for a certain species of animal to migrate in the summer, show all the public buildings where children are not allowed to enter, etc*)
- Availability of information at different levels of detail (*e.g.: can perform queries on different levels: 1) show every kindergarten in a certain area, every school, or every language teaching institution; 2) ask for all educational places for children under 18; 3) show all educational places; 4) show all public buildings; 5) show all buildings in a certain area*)
- Dynamic access to information that otherwise would be difficult to obtain [3] (*Using WS information in remote locations can be accessed over the Internet, by adding semantics this can be done automatically, without having to look manually through every single WS specification*)
- Avoiding the separation of data based only on their representations (*semantic meaning is added to the data, relationships among the entities and their behaviour are explicit declared*)
- Allowing the creation of ad hoc services that meet the users' interests. (*users can customise their needs and ask for broader and more customised services to be achieved*)
- Creating representations of incomplete information (*if working with several sources and some information is missing from one of them, it is possible to generalise them and work with a broader, but more complete data classification*)
- Permitting different levels of detail (*due to the Ontology classification, it is possible to go 'up and down' in the classification, depending on the level of detail needed*)
- Enabling multiple representations of geographic space (*using ontologies, several views of the same data are possible. It is also possible to translate from one data representation to another one automatically*)
- Solving heterogeneity problems (*data represented differently can be merged*) [5][6]

This use case will mainly focus in the last point: how to *solve heterogeneity problems* within location based systems. Ontologies can help to integrate various systems with different kinds of data and/or representation.

A GIS-domain ontology can be used, not only to integrate GIS, but also to integrate other kinds of documents; such as databases with relevant information or documents defined in natural language terms.

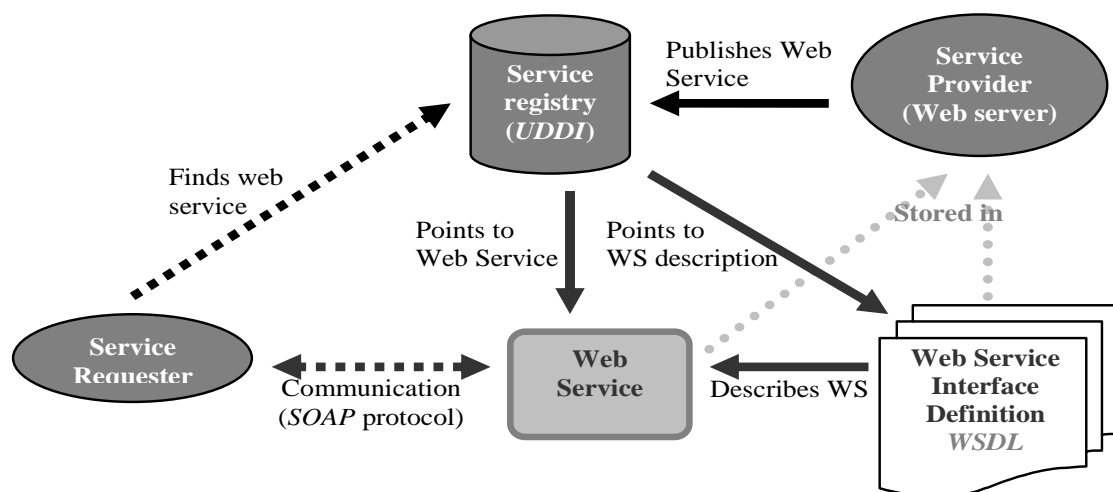
These systems are of enormous potential future value. Having an Ontology that describes a spatially-related data repository, makes it very much easier to co-operate with other systems. It is also of benefit in terms of data description. Attaching an Ontology to a system, makes it self-described and fully documented, so that users and developers (as well as automated programs) know the exact meaning of every field and the overall structure of the database (instead of having it documented in natural language, which can causes many ambiguities, or not documenting it at all, so that a user has to go through every single field of all the tables in the database to understand what they refer to) [7].

## 2.3 Building WS and SWS over a GIS

### 2.3.1 WS technology

WS are well defined, reusable, software components that perform specific, encapsulated tasks via standardised web-oriented mechanisms. They perform actions over the web, connecting several computers using the Internet to exchange data. They can be discovered, invoked and the composition of several services can be choreographed to get richer information.

WS technology is composed of several components: a WS interface (which is described in WSDL language) describes the service functionalities. This service description is registered in a service registry (UDDI), where the user can look for it. Once the desired service is found, it can be retrieved from the server where it is published (using an XML-based protocol: SOAP). The next picture depicts the process of publishing, finding, and consuming WS over the web:



- **Web Service:** A self-contained, modular application that can be described, published, located, and invoked over the Web. Modular applications that can perform any kind of request (no matter how simple or complex it is), which can be accessed across and via the Web.
- **Service requester** any human being or automated program that invokes the service.
- **Service provider** the respondent to requests, normally WS are stored and provided by Web Servers.
- **Service registry:** contains the description of the WS published by the Service providers. It is a way to advertise the WS so the Service requester can find them and invoke them afterwards. The registry holds a description of every service published in it, which can include a profile of the provider of the service (company name, address, business, etc), a profile of the service itself (name, category, etc ), and can also contain a pointer to the WS interface definition (WSDL description)
- **Web Service Interface Definition:** Usage interface. Describes the operations the service performs, in terms of **input and output messages**; and associates the messages to physical endpoints (network locations where the service can be invoked)

### 2.3.2 Benefits of WS over a common repository

There are some advantages in implementing a WS-based solution rather than creating a common spatial data repository like **ViewEssex**:

**ViewEssex** will collate data from different kinds of source, but the updating of data into the repository has to be supervised and controlled by a qualified user, who has to agree on the data and its format, before it can be uploaded into the server.

For this purpose a GIS forum has been created. There, GIS experts can decide what data is relevant to be uploaded. The forum also engages with the help from the police and fire services in the county. An IT company also gives advice on technical issues.

If instead of merging the systems the use case develops WS to wrap every system functionality these problems would be solved, as every repository could share their functionalities with anybody via the Internet.

Another drawback of a common repository is that if users still work with their local copy of the data, and they update it or change it, then the data in the common repository will be out-to-date. There has to be an agreement on how frequently the participants pass data into the common repository. Developing WS around the initial repositories instead would solve this inconsistency problem as every organization is responsible for their own WS and there is no need to maintain or upgrade additional datasets.

While ViewEssex adopts an open standards approach, the repository does entail certain dependencies. Vector data stored in Oracle Spatial can only be uploaded or read by software compatible with the Spatial geometry formats. Also, ECC's image data is stored separately, in proprietary raster formats like ECW, which are not supported by every possible tool. Therefore, complete harmonisation would involve a significant amount of manual work and decisions between parties, which could be easily avoided using WS.

But the simplest and more important advantage of WS is that there is always, somewhere, data to be merged. No system can hope to merge all the relevant data in the world (GIS, economic datasets, social datasets, weather, etc). There are always systems that hold interesting data for other systems somewhere in the world. Using WS, these data can be shared with anyone through the Internet, making it possible to improve the quality and accuracy of any system.

### **2.3.3 Benefits of SWS over WS**

Although WS provide several benefits to merge data spread out all over the Internet, they still lack of several functionalities.

It is a difficult task for the user to find the most suitable WS that solves his particular problem. The user has through all the published WS interface definitions (WSDL), to find the capabilities that best suit his purposes. The WSDL definition is not powerful enough to enable automatic management by machines as it defines the service capabilities only in terms of its inputs and outputs. Human interaction is essential to interpret these definitions.

On the contrary, SWS can be automatically discovered over the Internet. SWS are a new type of application. They are "*self-contained, self-describing, semantically marked-up software resources that can be published, discovered, composed and executed across the Web in a task driven automatic way*" [Arroyo et al. 2004]

SWS are a type of intelligent Software Agent (software agent capable of flexible autonomous action in order to meet its design objectives) which provides intelligent access to heterogeneous, distributed information over the SW. There has to be an unambiguous frame and a common language for SWS to operate, which is obtained through Ontologies. The Ontologies lead the conversation between SWS, stating a common domain of discourse. This way SWS can automatically accomplish tasks on the Web, "taking the human out of the loop".

SWS are considered the ultimate tool to change the way business and other services are delivered through the Web. They can be discovered, executed and combined by other automated software agents without any need of human guidance.

The main key benefit foreseen of using SWS within eGovernment case studies is to eliminate or minimise human operations (e.g. re-keying, re-purposing or re-formatting information sources) to enable correlation of different sources for decision-making,

statutory reports etc. In the use case, Ontologies and SWS will be applied to the GIS scenario to automate the integration of different data sources spread out across ECC, other partners and over the Internet.

For example:

We want to look in the internet for the *cheapest areas to build a house around London, and with good communication with other European countries. (Our job is based in London but it involves a lot of travel because we are working in an international company). Moreover, we want to have a garden at least 30 square metres and we don't want a noisy or polluted area as we have two children.* If working with WS the user would have to look through every single WS description, trying to guess what the services do based on their inputs and outputs. As this is a complex query, there might be no single service that accomplishes this task, so several WS have to be invoked. If doing it manually the user has to compose all the results himself, paying special attention to inputs/outputs.

If using SWS the integration task is much easier for the user, as it is fully automated. WS are automatically composed and choreographed. Following the WSMO [23] architecture (WS, Goals, Mediators and Ontologies) all the user needs to do is to invoke a goal (a goal defines the objectives that a client may have when consulting a WS. At least one ontology is imported by every goal to define its terms.). This goal would be decomposed in several sub-tasks, which are achieved by different services and automatically composed back into the main goal. For example, the goal stated above can find the following services:

- WS to locate all the land in the UK with permits to build houses
- WS to locate all the land in South-East UK with permits to build houses
- WS that list some parcels for sale in an Essex estate agent
- WS that provides a list of houses for sale in a Brentwood estate agent.
- WS that locates London, Essex and Brentwood in the South-East part of UK
- WS that locates all the airports within 200km from London
- WS that provides all the destinations of all the English airports
- WS that sorts the destinations into European – Non European cities.
- WS that locates the 10 nearer ports to London
- WS that locates all the train stations in England with European-destinations
- WS that provide real-time information about the highways around London, traffic congestion, accidents, and estimated times to get into the City.
- WS that provides the noise levels in south-England
- WS that state the maximum noise level a person can tolerate (above that level it is considered dangerous)
- WS that provides a list of the noise level a person can tolerate depending on his age (children and elderly tolerate less noise than young and adults)
- WS that provides pollution data in the Councils of London, Essex, Hertfordshire, Berkshire, Buckinghamshire, Surrey and Kent.

All these services are automatically discovered by the goal, using several ontologies to guide the search. As some of these services provide the same information differently, the most useful ones (depending on other user preferences) have to be

discovered and composed automatically (this is done with the help of the mediators). The results are sent back to the user, who does not need to be concerned about the searching or composition of the service.

## 2.4 GIS Scenarios

Four potential scenarios were considered at the very first stage of this project:

- **Emergency planning** (How to prevent, and what actions should be taken in case a flood, fire, explosion, etc occurs)
- **Traffic congestion** (Information about the traffic in a certain area of the County)
- **Stansted expansion** (What the consequences of the Essex airport expansion would be)
- **Housing scenario** (What the consequences of centrally planned development expansion in Essex would be)

These four scenarios were defined at the beginning of the requirements specification phase. All the scenarios were equally considered and discussed with the experts in the interviews the ECC-DIP team carried out.

When defining each scenario, it was realised that the two first scenarios (emergency planning and traffic congestion) would require special attention because they make use of real-time data (data that is constantly changing in time). Their behaviour is very different from both operational and management support systems [2] If one of these is finally chosen special attention would need to be paid to this issue. These two scenarios can be seen as a combination of static/slowly varying data (e.g. Evacuation route, road network ...) on top of which the real time data is superimposed. If real-time WS are deployed, then SWS can be easily created on top of them, forming a very powerful semantic real-time system. Housing and Airport expansion on the other hand, have a lower dependency on the use of real time data.

## 2.5 Identifying key experts

The next step, after identifying possible scenarios, is to discover a set of key users who have the credibility to define the software to be built. They have been interviewed to provide guidance in defining software business requirements. These users were carefully selected by the ECC development team, based on their knowledge about GIS and the four pre-selected scenarios.

## 2.6 Interviews

Interviewing is almost certainly the most widely used fact finding method. They are used to acquire knowledge about the current system and about user requirements [2].

### 2.6.1 Description of interviews

The WP9-ECC team held a total of five meetings, in which eight relevant people were interviewed. The experts were grouped within their areas of expertise so that it was more convenient to conduct the interviews. Most of these interviews lasted about 1 hour.

During the interviews with the abovementioned a better understanding was obtained of how geographic information is currently managed within ECC and how data is shared between the ECC and other bodies (and what data the ECC would like to share and with whom in the future). The other main purpose of the interviews was to survey the experts to find out which scenario would be most relevant for their daily work in the ECC.

When commencing the interviews, a brief introduction about DIP and SWS-enhanced systems was given to provide the interviewees with a better understanding of the project. The interviews were guided by a list of questions previously developed by DIP project staff in ECC and NIWA. Very useful assistance was also provided by an ECC GIS-expert. The interview questions can be consulted at Annex 2.

### 2.6.2 Overview of interview results

While carrying out the interviews the ECC-DIP team took notes from the answers given by the experts. This information has been formatted in Table 1 below. It shows the main issues covered by the different scenarios (Stansted expansion, housing, traffic congestion and emergency planning). Crime and disorder is a further scenario that emerged from the interviews.

The information gathered from the interviewees has been clustered within several relevant topics: main goals of each group, the kind of people the system would focus on, the different systems involved in each scenario, their data sources and the kind of data managed in each scenario (pointing out whether they make use of real-time data), how the data is displayed and the main problems which these systems currently face.

Table 1 reflects the information gathered from all the interviewees, grouped by different areas of interest. The yellow-coloured cells highlight the issues the prototype will focus upon. It is noticeable from the table that scenarios overlap at some points.

Stansted	Housing	Traffic congestion	Crime and disorder	Emergency planning
<b>Main goals</b>				
Model and forecast environmental impact (e.g. noise levels) in relation e.g. to runway positioning	Track and analyse planning permissions	Reduce journey times	Improve analysis and forecasting of crime and disorder	Risk assessment in relation to specific geographic areas (e.g. plane crashes, floods, explosions, terrorism etc)
Model and forecast socio-economic impact	Analyse and forecast demographic trends (impact on housing)	Produce statistical data to assist forecasting		Response planning e.g. evacuation
Model and forecast economic viability of airport (income versus costs)	Spatial analysis			Make data available for use by other agencies locally, nationally
Model and forecast compensation requirements	Model and forecast impact of house building on traffic/transportation			
	Forecast impact on housing of Stansted airport expansion			
<b>Target audience</b>				
Politicians, planners	Politicians, housing planners, schools service, police	Politicians, traffic planners, public	Politician, planners, law enforcement agencies	Politicians, emergency planning agencies, public

Stansted	Housing	Traffic congestion	Crime and disorder	Emergency planning
<b>Systems involved</b>				
	GIS-related database	Traffic Control Centre database COMET (Siemens) – triggers decisions based on available info – in development	(Separate system for Crime and Disorder – Manchester U)	
<b>Data types</b>				
GIS	Lists of sites (numbers and types of dwellings) pending permission	GPS positioning systems/speed calculations	Crimes	Chemicals, poisons
Archaeological	Digital boundaries (GIS)	Traffic flow counters	Victims	Weather
Environmental impact		Onboard data from buses and other vehicles	Socio-economic	Flooding
Socio-economic		Weather		Historic data (e.g. plane crashes)
		Events (e.g. V Festival)		
<b>Data Sources</b>				
Will use <i>ViewEssex</i> soon	Currently using <i>ViewEssex</i>	Might use <i>ViewEssex</i> soon	Police	Currently using <i>ViewEssex</i>
Stansted Airport	University (outsourced demographic forecasting)	Private sector (GPS): Traffic Master IT IS: data from heavy goods vehicles	Youth Offending Teams	Stansted Airport

Stansted	Housing	Traffic congestion	Crime and disorder	Emergency planning
<b>Data Sources</b>				
Airlines (e.g. British Airways)	Consultants are employed to merge data and produce studies	Bus Operators	Victim Support	Emergency services. Police, fire, ambulance
Environment Agency	District councils: land and property gazetteers	Police	Ambulance Service	Department of Health – Poisons Unit
District Councils (e.g. Uttlesford)	ODPM/IDeA (soon?)	Highways Agency		Many different publicly available sources
		Internet		
		Radio station		
		Public (phone calls)		
<b>Real time data?</b>				
No	No	Yes	No	Yes
		No		No
<b>Display/ presentation</b>				
Analysis shown on individual maps: no layering	GIS (on maps)	Analysis already shown on maps with manual layering (in Traffic Control Centre)	ArcView (GIS)	ArcView (GIS)
	Spreadsheets	Web site soon to be launched		Website with a public and private section
				Database, CDs
<b>Problem areas</b>				
Merging heterogeneous data (currently done manually)	Different data systems, different standards	Main roads only covered by most data systems	Heterogeneity (applying spatial co-ordinates)	Heterogeneity of meaning (international words, spatial co-ordinates)

Stansted	Housing	Traffic congestion	Crime and disorder	Emergency planning
<b>Problem areas</b>				
Some data types unavailable			Merging heterogeneous data (currently done manually)	
<b>Future plans</b>				
	Move into forecasting	Forecast impact of housing expansion, Stansted Airport etc	Secure web	Better map-based outputs: ease of evacuation planning
	Make data accessible to users electronically (currently done manually)	Accessibility planning	Modelling and forecasting (ad hoc relationships)	Improved modelling and forecast 'what-ifs) – DSS.
	Merge data on the web	Personalised traffic control systems	Provide different user views on data	
<b>Time critical issues</b>				
2 <sup>nd</sup> runway decision first half 2006				

**Table 1: Summary table of interviews results**

## 2.7 Scenario selection

By carefully studying the commentaries obtained from the interviews and the results of the analysis of some of the systems to be merged (mainly **ViewEssex**, [1]) the ECC-DIP group has chosen the final scenario. The table shown in Table 1 has been a useful instrument in making this decision. Key issues have been highlighted there.

The main issue in a spatially-related data project is the place where things happen. So the first need is to define a common area to focus on. As Stansted Expansion has proved to be a key issue for ECC and provides flexibility in terms of accessible datasets and people willing to collaborate, it was decided to base the use case scenario around the Stansted area.

The picture below (Figure 1) depicts where Stansted airport is placed in England, and its situation in relation to Chelmsford and London.

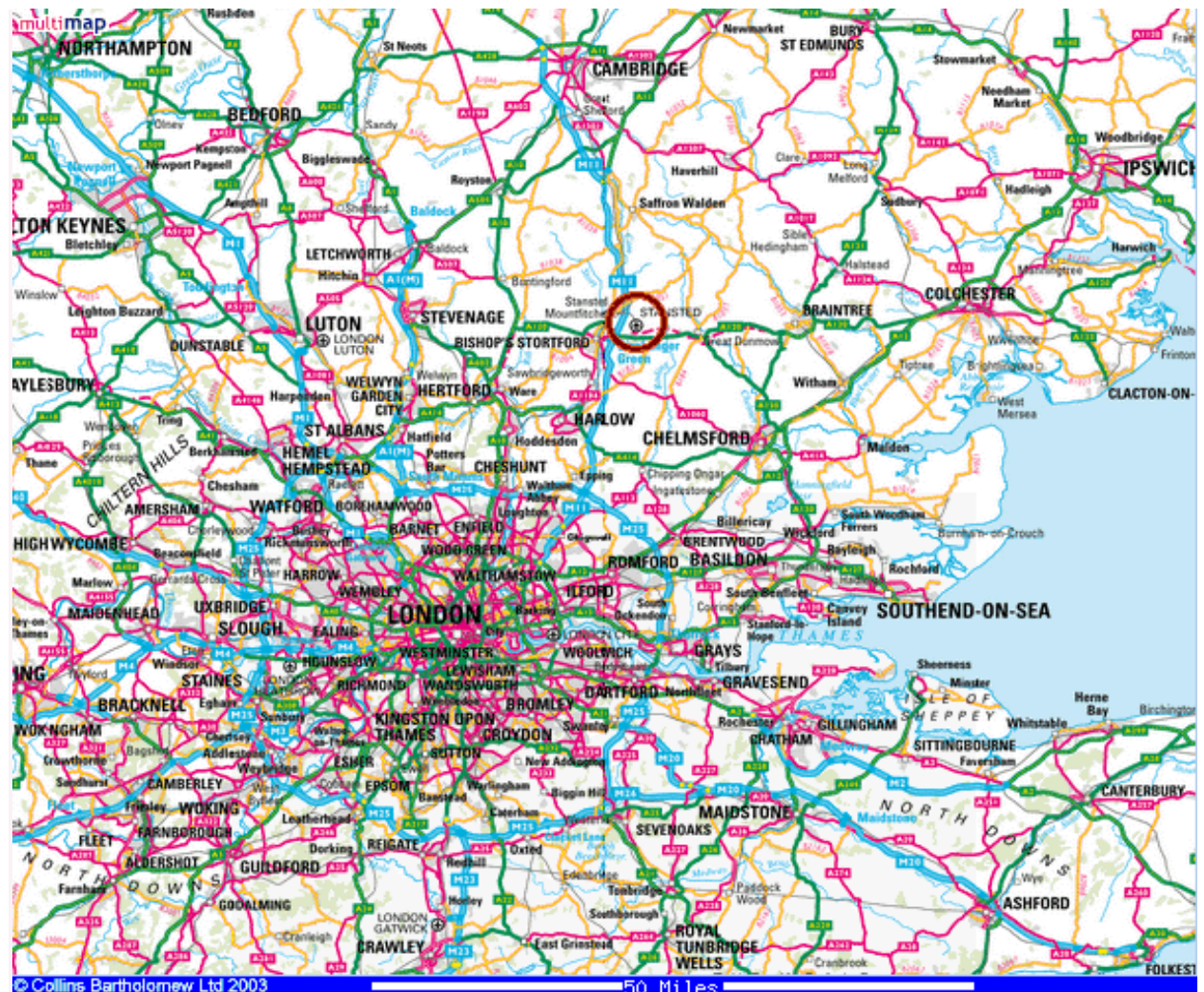


Figure 1: Stansted map. Source: [www.multimap.com](http://www.multimap.com)

Next picture (Figure 2) shows a closer view of the Stansted airport and its surroundings. This map has been obtained querying `ViewEssex`:

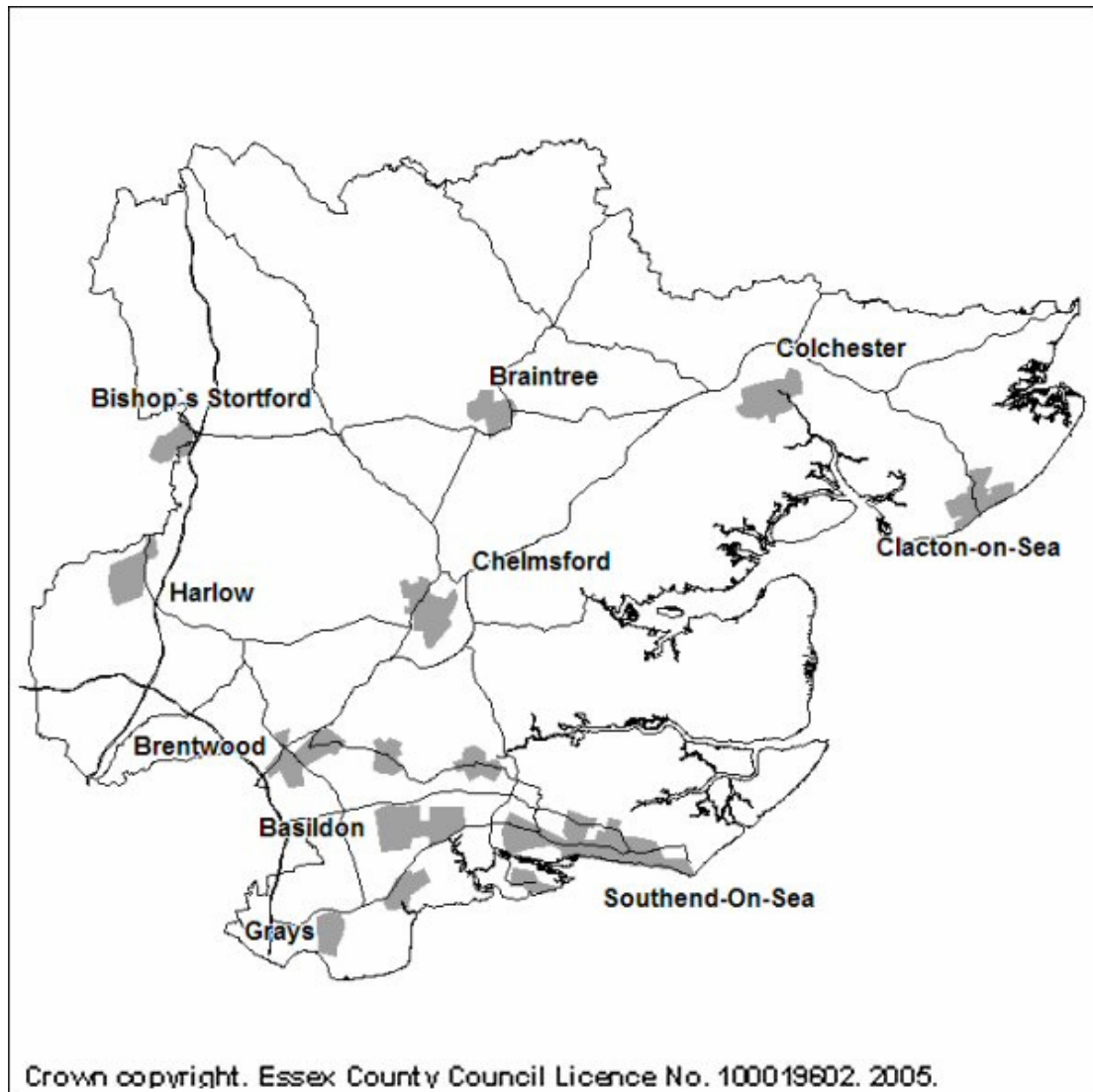


Figure 2: Stansted map. Source: ViewEssex

Once the area is defined, major scenario issues need to be selected. It has been decided to concentrate on the effects of the proposed Stansted second runway expansion.

The prototype will be focused on the needs of politicians and decision-makers, rather than for use of the general public. It will support decision-makers to assess the effects of placing a second runway in a certain area. The system will help evaluate the impact and respond to certain economic, sociological, and environmental questions. The system will support any decision with figures and data, relating all the information that is currently distributed across many different data sets belonging to different organisations.

Another issue that is going to be targeted through the use case prototype is the housing scenario. This will be addressed within the boundaries of the geographical area around Stansted. The housing scenario is closely related to the airport expansion, which implies a greater need to build more houses and/or hotels around it. The system to be created can help measuring the impact on the environment, roads/traffic, social issues, security, etc. Figure 3 is a map of Essex County, depicting the main urban areas:



**Figure 3: Essex map. Source: ViewEssex**

In addition to these two main issues, attention will be paid to related issues from the other scenarios, such as: what is the most likely area for a plane-accident to happen? (Emergency planning scenario), forecasting how the airport or the housing expansion will affect congestion on traffic and trains (traffic congestion scenario).

We are only going to focus on planning and prediction, not in response systems as these make use of real-time data, which would be more difficult to obtain. Also real-time systems need a very high time-response, which is difficult to obtain in a prototype. For these reasons, natural disasters, evacuation routes, chemical disaster response... are out of the scope of this prototype.

## 2.8 Potential systems to be merged

Based on the business case for the **ViewEssex** project, and the interviews, several data sources have been identified, the data from which ECC consider can beneficially be merged, including:

- District Councils/UAs (Unitary Authorities)
- Fire/Police
- Subcontractors
- Ordnance Survey

It is also important for the ECC to share data with:

- District Councils
- Emergency services
- Central Government

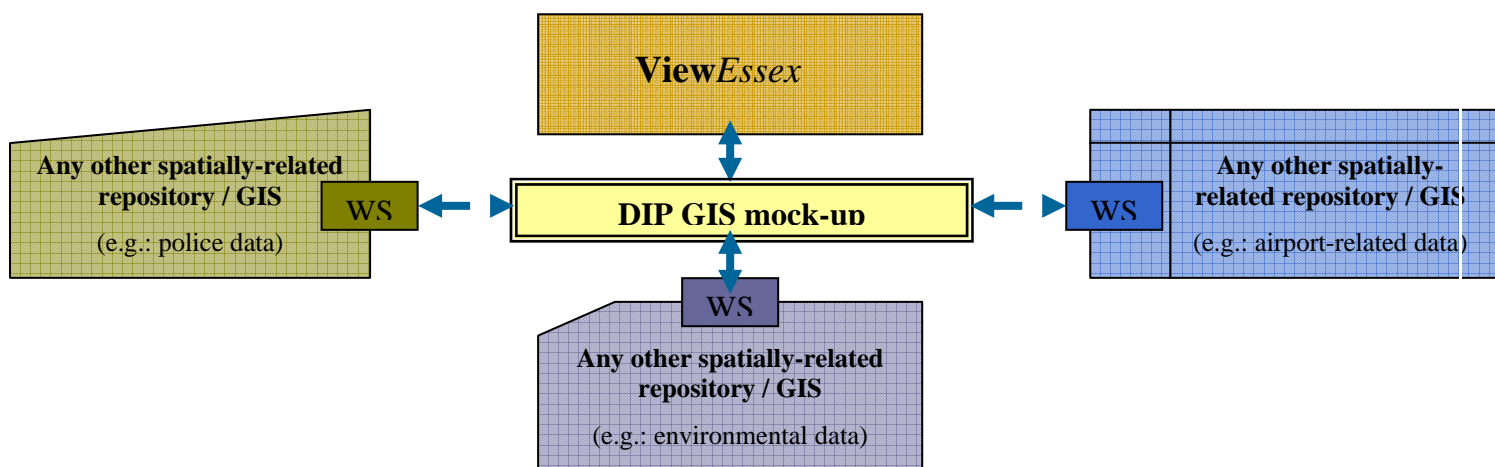
In a future series of interviews, the aim will be to identify the necessary data sets, including the above organizations where possible. The use case prototype will seek to merge the maximum number of these systems using Semantic Web Services technology. The number successfully finally merged depends on time constraints and the eventual legacy impairments we encounter in gaining access to the data.

**ViewEssex** will be the core dataset, as it contains a large quantity of useful data and already belongs to ECC, allowing free and unlimited access, as long as its normal functions are not affected. ECC and the developers of **ViewEssex**, are interested in **Web Services** deployment in the future. As they state in the GIS Business Case (Appendix 15) “*Web services can be accessed from a wide spectrum of software products and – as well as offering a means for sharing map data with external bodies such as the Districts/UAs, Fire/Police and sub-contractors*” [1]. Currently there are no WS deployed over the **ViewEssex** project, thus the DIP project team in Essex will need to develop them. However this will allow subsequent re-used in future phases of **ViewEssex**, resulting in cost savings for ECC.

### 3 PROJECT CONSTRAINTS

#### 3.1 Implementation Environment of the Current System

The current system is composed of several isolated systems all over the ECC and other organizations (Police, Airport, Hospital, Fire Station, etc), but **ViewEssex** is going to integrate a large part of these ECC GIS. Next picture shows how DIP mock-up fits into ECC GIS:



**Figure 4: How DIP project fits into ViewEssex and mediates with other projects**

The DIP prototype will make use of **ViewEssex** Web Services, if there are any deployed at the time of implementing the system. Otherwise new WS will be created, which can later be reused for **ViewEssex** purposes.

Besides **ViewEssex** many other data sources are still isolated. The main purpose of this prototype is to demonstrate how SWS makes it possible to easily merge GIS and other kind of spatially-related data. The nature of the datasets may be different and heterogeneous, but SWS can easily handle this.

#### 3.2 Ontology Constraints

The main purpose of the Ontology/ies to be built is to describe the elements (terms, vocabulary), their relationships and the rules of the given domain, and to describe the knowledge in a generic way, so everybody committing to the same Ontology has the same understanding of that domain.

The GIS Ontology (domain ontology of the GIS scenario) will be the backbone of all the systems to be developed. It will provide the basis for the rest of the WSMO [23] elements: they will provide the semantic description for: the service capabilities, as well as describe the goals and mediators. In this way WS are semantically described, and can be discovered and invoked by other services automatically.

Different Ontologies will be provided (geographic coordinates, legacy issues, pollution, economic data, etc...), one for each system to be merged. Each Ontology will describe the domain of each system to be merged. Some of them may overlap or contain the same issues expressed in different ways (an Ontology reflects the reality of the system it is focused on). As there will be several Ontologies (presumably one for each system to be merged), it will be necessary to make use of WSMO [23] mediators (more concretely, OO-mediators) to integrate them all. An OO-mediator is a special kind of WSMO [23] component that helps to map between different Ontologies.

The domain Ontologies will be developed in ECC .The Ontologies will be developed in OCML but can be easily translated into WSML [24] at a later stage. The Ontologies will be imported in IRS-III and used as the basis for all WSMO [23] elements: goals, mediators and SWS descriptions.

The first prototype is planned to be implemented in IRS-III (WSMO [23] compliant), but the second prototype will be probably implemented in WSMX [25], and make use of the DIP API WSMO4j [26].

### 3.3 WSMO operations required

This scenario is focused on data integration. The use case will demonstrate how SWS can improve traditional methods and automatically integrate different kinds of data distributed across different sources.

For this purpose it will be necessary to make use of the following WSMO [23] features:

- **Composition:** Choreography of several WS which together can achieve the invoked goal. It is probable that no WS will completely fulfil the selected goals, and it will therefore be necessary to compose several WS that achieve that goal partially.
- **Mediation:** Several types of mediators may be needed to mediate between different WSMO [23] components: oo-mediators (OntologyOntology-mediators) to mediate between different types of data, wg-mediators (WS-

Goal mediators) to link the goals with the services, ww-mediators (WS-WS mediators) to link two services and gg-mediators (Goal-Goal mediators) mediators that link two goals.

- **Discovery:** It is necessary to discover the best WS that fulfils the defined requirements, which are expressed by the user in the form of a goal. This is called: domain-dependant discovery. The user does not need to know the implementation of the WS, its inputs, outputs or any other features, only to invoke a goal, selecting it from a goal repository (stored in the IRS-III in this particular case)

## 4 REQUIREMENTS

Once the interviews have been finalised, the next step is to identify the requirements of the new system. The information gathered by the ECC-DIP group in the interviews is studied and split up into one of the following categories: functional requirements, and non-functional requirements.

### 4.1 Functional Requirements

This section defines the functional requirements for the SWS-based prototype that ECC is going to create for the GIS scenario. Functional requirements (often referred to as functionality) capture the intended behaviour of the system. They describe what a system does or is expected to do. This behaviour may be expressed as services, tasks or functions the system is required to perform: [2]

- Descriptions of the processing that the system will be required to carry out
- Details of the inputs into the system from paper forms and documents, from interactions between people, such as telephone calls, and from other systems
- Details of the outputs that are expected from the system in the form of printed documents and reports, screen displays and transfers to other systems
- Details of data that must be held in the system.

#### 4.1.1 Objectives

The main objective of the prototype is to demonstrate how ODGIS (Ontology-driven GIS) can obtain better results than plain GIS. It will be demonstrated how SWS based systems obtain a better performance than systems based on manual data integration. Coping with data heterogeneity is the main advantage of these kinds of systems. That is the main purpose of this prototype: to show how SWS-based systems can automatically integrate different systems and data sources into one system, improving the functionality of the separate systems

### 4.1.2 GIS systems integration

As already stated in previous sections, one of the main goals of this prototype is to integrate data from heterogeneous sources. This leads to the following functional requirements:

- [FR 1] Integration of several data sources.** The prototype will integrate several GIS and other kind or spatially-related data. One of these systems will probably be the **ViewEssex** repository, being developed by ECC. The others are still to be chosen, but there could include: police records, airport or air company databases, pollution and noise statistics, weather, crime statistics, land registries, etc.
  
- [FR 2] Real world databases and data.** The prototype should operate with real world data, in order to reveal its full potential. Access to real data is probable, but in case problems arise (legacy, data privacy, protection...) a reconstruction of the data will be created, as accurately as possible.
  
- [FR 3] Independent test system.** The prototype should be implemented on the real world installations, or if this is impossible, on the most accurate test systems.
  
- [FR 4] Anonymised data.** As this system is simply a prototype, no security will be implemented. Consequently any personal data contained in the integrated data sets must be anonymised. If no anonymised data is accessible a dummy database should be created. If this prototype is successful and it is to be implemented in the real world, trust issues will be taken into account at that stage.
  
- [FR 5] Web Services.** As this prototype builds SWS on top of ordinary WS, these have to be provided by the datasets to be integrated. If the datasets do not provide this service, they will be implemented on top of them, encapsulating the data. There are various formats in which IRS-III can obtain and use these inputs (proper WS, java code or lisp code)

### 4.1.3 WS Requirements

As stated in [FR 5] above, several WS have to be created to wrap the datasets in order to serve as an input for the SWS. Below are listed all of these services and their functionality:

- [FR 6] **WS do not interfere with data.** All the WS to be deployed can only access the data repositories. They will not be allowed to delete it or modify it by any means.
- [FR 7] **WS provide basic maps.** These WS will provide basic maps of the targeted area to the user. On top of these maps the user can make inputs and will receive the depiction of the outputs. These maps can be of various types: (roads, rivers, forests, population, towns, etc) and formats (raster or vector data)
- [FR 8] **Advanced WS.** These services will provide more customised information, such as the number of people living in a certain area, number of children attending a certain school, number of people using a certain hospital, etc.
- [FR 9] **Other WS:** As access to the data sources is obtained, it will be possible to define other services on top of them to access that data.

#### 4.1.4 Special WS Requirements

- [FR 10] **Spatial extensions:** Support for operating on spatial data e.g. support for querying areas and geometric shapes: point X is within 500meters of point Y. Basic SWS can be invoked by other SWS.
- [FR 11] **Geocoding of addresses:** There should be one or more SWS that perform this operation. This refers to relating non-spatial data (e.g. a telephone directory listing) to spatial data (e.g. a GIS) via coordinates. First the non-spatial data sources have to undergo a procedure to obtain those coordinates. These SWS would be called by many other upper-level SWS.
- [FR 12] **Reverse Geocoding WS:** There should be one or more SWS to perform this operation, which enables addresses/points of interests within a given geospatial shape to be identified (i.e. transforms coordinates areas into addresses)

#### 4.1.5 Graphical User Interface Requirements

The prototype is Web-based. This method has many advantages, such as reducing drastically the amount of software the user has to install, and allowing distributed responses over the Internet, as well as concurrency (several users working on the same data at the same time).

- [FR 13] **Web-based User Interface.** The prototype must be accessed from a standard web browser. The communication with the databases will be done through the Internet, so the user does not have to install any copy of the data onto their computer.
- [FR 14] **Graphical-textual interface.** The interface will allow the user to create goals, execute them and have the results displayed afterwards. The goals can be expressed in two ways: graphically (a simple click on a map, selecting a point/area), or choosing an option from a list.
- [FR 15] **Results web-page.** The results of the goal execution have to be presented back to the user. Some of them will be displayed as overlapping layers on the maps. Others will be statistics or a list of figures in plain text. These pages are dynamically created depending on what SWS have been executed.
- [FR 16] **Goal invocation.** As already explained, the user can invoke goals from the user interface. In any case, the formulation of the goals should be hidden as far as possible from the user, who will not be aware of what SWS are invoked. This is a completely transparent process for the user
- [FR 17] **Goal definition schema**

The following schema shows the functionality of the system, defined by the goals that can be invoked by the user. Sub-goals are goals that depend on other goals to be executed beforehand. These goals may not be definitive, as different data sources are still being surveyed. This schema therefore is intended to be an example of goal description and sub-goals decomposition rather than a definitive design.

The goals are defined here in plain English but will be translated into WSMO [23] when implementing the system. The goals will then be linked to the services that fulfil them, and may make use of several mediators to cope with heterogeneity problems.

#### **Goal 1. Stansted expansion**

**Goal-definition.** The user draws an area where he wants to place the second runway.

**System response.** Draws the area selected on the map

##### **Sub-Goal 1.1. Private/Public buildings demolished**

Show houses/schools, offices, industry, hospitals, fire-stations/ police/stations (etc) to be demolished.

**Goal-definition.** User selects this option from a list of options

**System response.** Shows on the map the elements affected and/or lists the buildings and their addresses in plain text.

**Sub-Goal 1.1.1. People affected by buildings demolition**

Shows how many people are affected by the property demolition.

**Goal-definition.** User selects this option from a list of options

**System response.** Lists the number of people affected in plain text (can show their age if information available)

**Sub-Goal 1.1.1. Cost of expropriating the buildings**

Estimates the cost of the expropriation

**Goal-definition.** User selects this option from a list of options

**System response.** Estimates how the government/companies would have to pay if they expropriate the buildings in the selected area

**Sub-Goal 1.2. Noise area**

Shows area affected by the noise

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the noise levels on the map (graphical output)

**Sub-Goal 1.2.1. People affected by the noise**

Shows people affected by the noise

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the people living/going to school /working/ hospitalized within the 'noise-area'. Graphical or textual output.

**Sub-Goal 1.2.2. Noise consequences**

Lists the effects the noise rates can have on people's health in the short/medium/long period

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the how the noise might affect to people's health. Textual output.

**Sub-Goal 1.3. Pollution area**

Shows pollution area around the airport

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the pollution levels on the map (graphical output), eventually stating the dangerous (unhealthy) levels.

#### **Sub-Goal 1.3.1. People affected by the pollution**

Shows people affected by the pollution

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the people living/going to school /working/ hospitalized within the ‘pollution-area’. Possibly outline people with asthma or other respiratory diseases. Graphical or textual output.

#### **Sub-Goal 1.3.2. Pollution consequences**

Lists the effects the pollution rates can have on people’s health in the short/medium/long period

**Goal-definition.** User selects this option from a list of options

**System response.** Shows the how the pollution might affect to people’s health. Textual output.

#### **Sub-Goal 1.4. Air traffic increase**

Estimate the number of new flights per day/week/month/year.

**Goal-definition.** User selects this option from a list of options

**System response.** Textual output. Makes a prediction of the flight number/frequency increase based on the length of the new runway.

##### **Sub-Goal 1.4.1 Travellers increase**

Estimate the number of passengers travelling from/to that area per day/week/month/year.

**Goal-definition.** User selects this option from a list of options

**System response.** Textual output. Makes a prediction of the number of passengers, based on the results of the “air traffic increase” goal.

##### **Sub-Goal 1.4.1.1. Traffic congestion**

Estimates how the increasing number of passengers will impact on the traffic congestion (private cars and public transport: buses and trains)

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical and textual output: number of cars/hour estimated in every road/ railway track.

**Sub-Goal 1.4.1.2. Parking problem**

As there is not enough parking space in Stansted people are increasingly using the surrounding provision for local inhabitants, then getting local transport to the airport. Estimate how many people would do that with the airport expansion and how it would affect the villages surrounding.

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical and textual output: Parking places occupied by travellers and how many people in the villages would be affected.

**Sub-Goal 1.5. Historical/Cultural impact**

Shows the historical buildings/cultural heritage sites affected by the airport expansion.

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical output. Highlights these sites on the map.

**Sub-Goal 1.6. Environmental impact**

Shows the green areas affected by the airport expansion.

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical output. Highlights these areas/buildings on the map.

**Sub-Goal 1.6.1. Environmental consequences**

Shows the species (animals/plants) affected by the destruction of the green areas.

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical and/or textual output.

**Sub-Goal 1.7. Facilities impact**

Shows the facilities that will be destroyed because of the airport expansion. Telephone lines, electrical installation, water supply, etc.

**Goal-definition.** User selects this option from a list of options

**System response.** Graphical output. Highlights where the facilities are located on the map and shows the area affected by the demolition (houses/public buildings, etc)

**Goal 2. Housing planning**

**Goal-definition.** User selects an area where they want to simulate housing building, and the number of new dwellings.

**System response.** Draws the area affected.

**Restrictions.** New houses cannot normally be placed over old houses/public building, although quite a lot of property development exploits previously used land. There are initiatives to encourage re-use of “brownfield” sites rather than rural land. Legal issues need to be considered.

Common goals: Goals that can be invoked after invoking any (or both) of the two main goals: Stansted and housing expansion.

**Environmental impact****Environmental consequences**

**Increase in number of people travelling** (referring to the new people commuting everyday)

**Traffic congestion**

Independent goals: Goals that can be invoked without invoking any of the two main goals: Stansted and housing expansion.

**Goal 3. Plane crash estimation**

**Goal-definition.** User selects this option from a list of options

**System response.** The system predicts where is the most likely place for a plane to have an accident, and the probability of this happening. Gives some statistics and shows former plane crashes as well. This goal can be invoked on the current airport or after the user expands it as desired.

If economic data is available, the costs of some of these changes could be estimated.

**4.1.6 Use case diagrams**

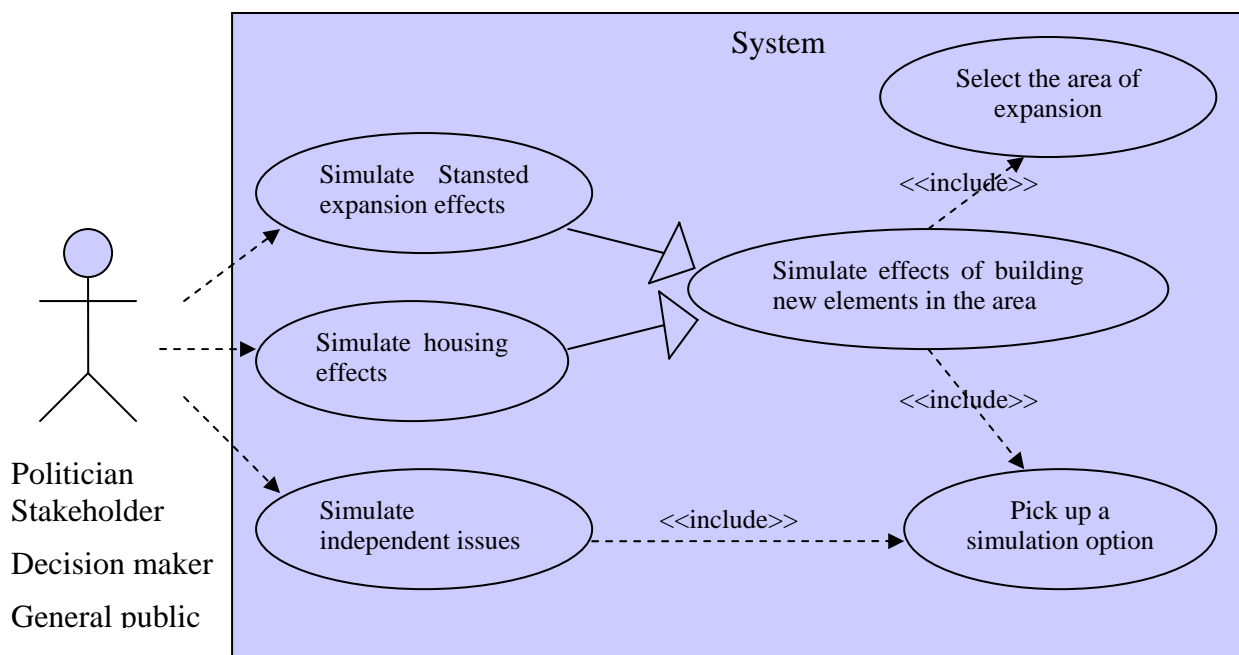
Use case diagrams are a technique to gather the user’s functional requirements in a high level overview of how the new system will be used. They are used to model functional requirements, because they focus on the functionality of the system, but are not as useful for modelling non-functional requirements. They are used to show the functionality of the system and what users will communicate with the system.

Functional requirements already listed in previous section describe the rules that guide the behaviour the system is going to have. On the other hand, use cases describe

units of this behaviour. One requirement may be in more than one use case, as one use case may make use of more than one requirement

*What is it that the users require. The use case describes the interaction as the user sees it, not as a definition of the internal processes within the system, or some kind of program specification. Use cases are supported by **behaviour specifications**, which specify the behaviour of each use case either using UML diagrams, such as collaboration diagrams or sequence diagrams, or in text form as use case descriptions [2]*

Figure 5 depicts the use case diagram of the whole system to be developed. It shows three main simulations: Stansted airport expansion consequences and housing consequences (which can be generalized in the more general use case: “simulate effects of building new elements in the area”) and other simulations/predictions about the current status of the area (which do not depend on any expansion)



**Figure 5: Use case diagram**

As depicted in the use case diagram, there are three main use cases: *Simulate Stansted expansion effects*, *simulate housing effects* and *simulate independent issues*. The first two can be generalised into the generic use case: *simulate effects of building new elements in the area*, which helps to clarify the picture. This generic use case always includes two operations: *select an area of expansion*, and *choose a simulation option from a given list of options*. The independent use case does not need to select an area of expansion, as it is independent of any building options; it only includes *picking up a simulation option* from a list.

### ❖ Simulate Stansted expansion effects

#### Use case description:

These use cases can be broken down into a step-by-step list of user's actions, and the system response. Next table shows in detail how the actor interacts with the system:

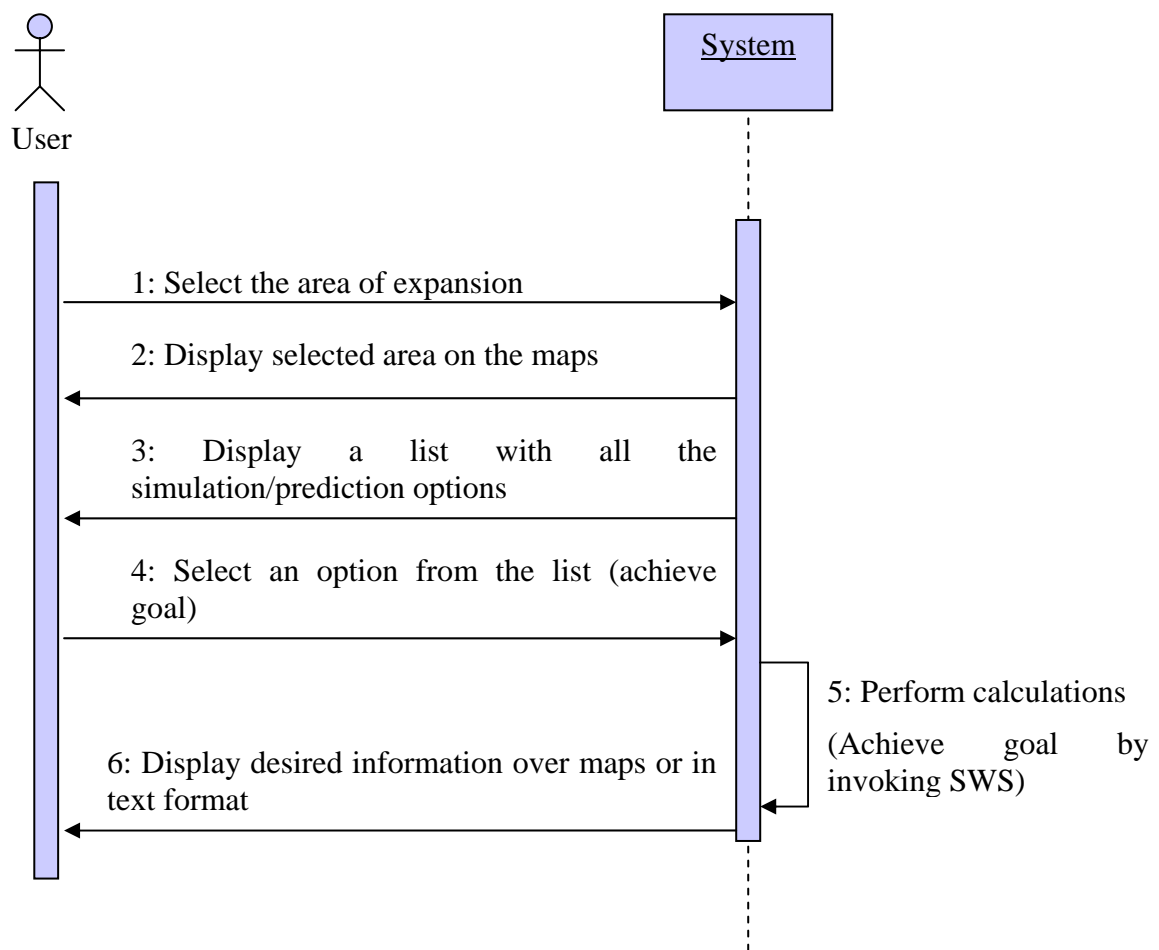
**Table 2: Stansted expansion use case description table**

<b>Actor Action</b>	<b>System response</b>
1. The actor selects the area of the expansion	2. Draws the area selected by the user
3. None	4. Displays a list with all the simulation options
5. Selects a simulation/prediction option from the list	6. Draws or lists the results of the simulation

#### High level sequence diagram

An interaction diagram is defined in the context of collaboration. It specifies the communication patterns between the user and several parts of the system. It contains a set of partially ordered messages (messages are placed in a time sequence, but two or more messages can be sent concurrently (at the same time)), which specify a communication (a signal to be sent or an operation to be invoked) as well as the roles to be played by the sender and the receiver. [2]

Iteration sequence diagrams show in a very clear way; how the interactions between objects are, and the sequential order in which they occur.



**Figure 6: Sequence diagram of the 'airport expansion consequences' use case**

### ❖ Simulate housing expansion

Use case description:

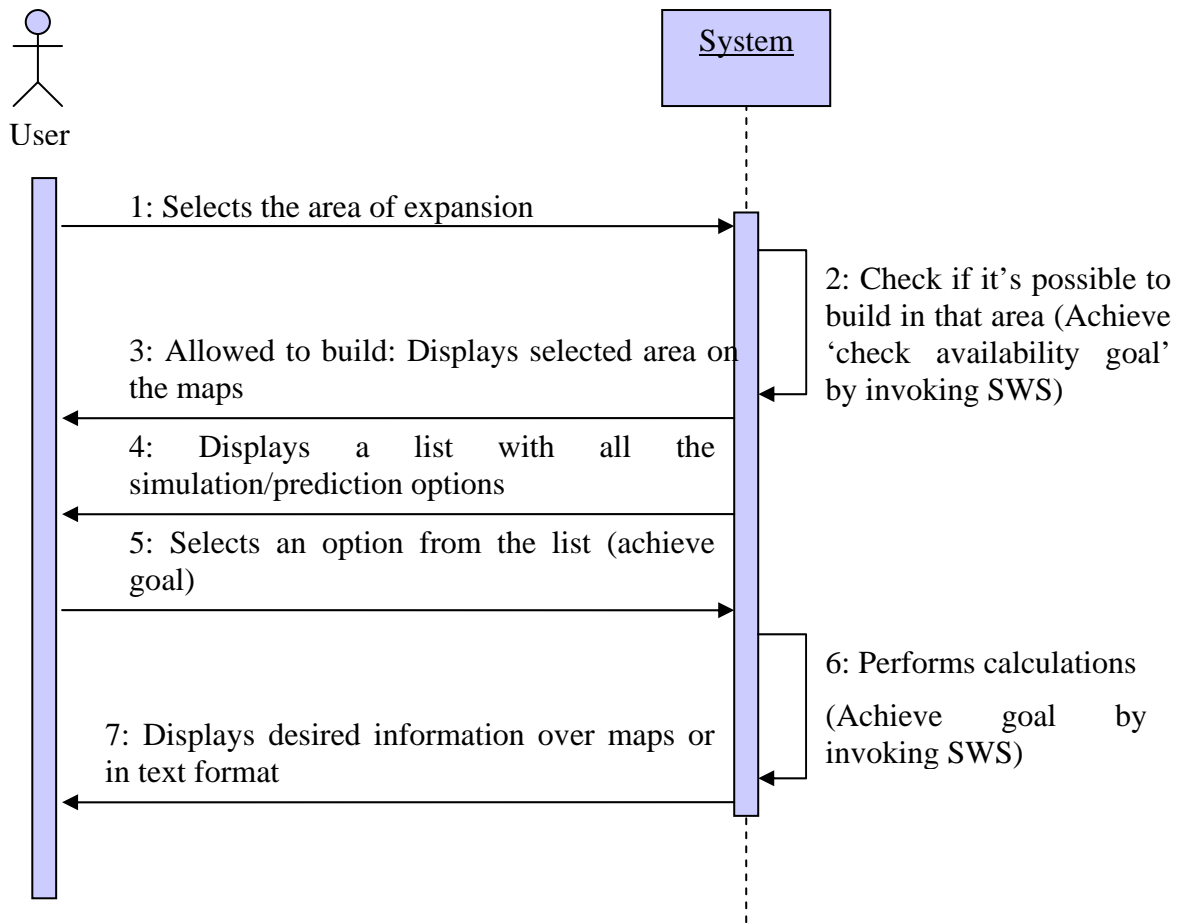
**Table 3: Housing expansion use case description table**

Actor Action	System response
1. The actor selects the area of the expansion	2. Draws the area selected by the user
3. None	4. Displays a list with all the simulation options
5. Selects a simulation/prediction option from the list	6. Draws or lists the results of the simulation

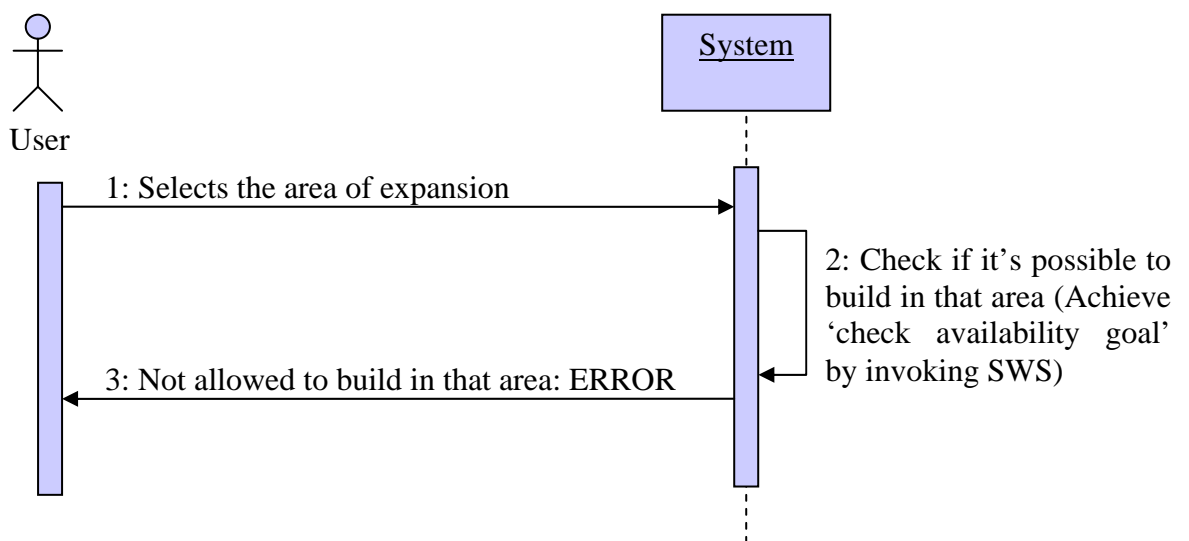
Alternative courses:

Step 2. The user is warned that in the selected area is not possible to build anything, and highlights the reason (legacy, property policy, previous houses or other constructions, etc)

High level sequence diagrams:



**Figure 7: Sequence diagram of the 'housing expansion consequences' use case – A**



**Figure 8: Sequence diagram of the 'housing expansion consequences' use case - B**

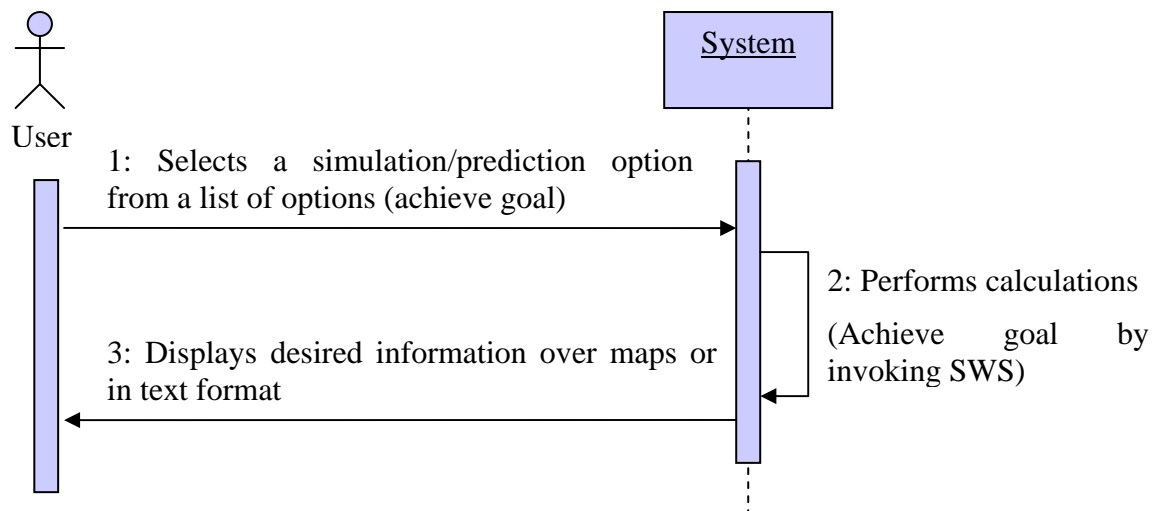
### ❖ Simulate independent issues

Use case description:

**Table 4: Expansion-independent use case description table**

Actor Action	System response
1. None	2. Displays a list with all the simulation options
3. Selects a simulation/prediction option from the list	4. Draws or lists the results of the simulation

High level sequence diagram:



**Figure 9: Sequence diagram of the 'simulate independent issues' use case**

## 4.2 Non-Functional Requirements

Non-Functional Requirements in Software Engineering present a systematic and pragmatic approach to 'building quality into' software systems. They describe aspects of the system that are concerned with how well it provides the functional requirements. These include the following 'quality' attributes: [2]

- Performance criteria: desired response times for updating data in the system or retrieving data from the system
- Accuracy
- Modifiability and scalability
- Security considerations

The next section lists the minimum non-functional requirements that have to be achieved in any application; although some of them might not be relevant for the use case, which is simply implementing a prototype to illustrate how SWS can enhance GIS.

#### **4.2.1 Accuracy**

The use case system is meant to be a demonstration prototype and not a working solution. However, it is intended to operate with real world data to demonstrate to the target audience the full potential of SWS. If any 'legacy system' problems arose it would become necessary to use test databases populated with anonymised dummy data, but these would not show all the features of SWS-enhanced systems to their full potential.

#### **4.2.2 Integration into existing infrastructure**

It is not intended in any way to interfere with the current GIS in ECC. The aim is to make a parallel case study to demonstrate how SWS-enhanced systems can collaborate with the current systems in improving ways of accessing heterogeneous data and merging different systems distributed across one or more an organisation.

Because most systems available in the different agencies involved do not currently provide standard WS interfaces to their functionality, it will not be possible to access the legacy systems without at least adding a WS interface to them. Some of the services developed for this prototype can be reused or serve as a basis for this future project.

**[NFR 1] Integration.** The SWS-enhanced prototype must not interfere in any way with the current work being carried out by the ECC or any of the other systems accessed. Data will be retrieved from these systems, but without altering it in any way.

As one of the main SWS capabilities is to integrate heterogeneous data across different systems and organizations, an important non-functional requirement is:

**[NFR 2] Distributed System.** The prototype will be implemented as a distributed system. Various components of the system will be implemented on different platforms and data accessed through WS will be spread across ECC and other external organisations

### **4.2.3 Scalability and Performance**

As SWS is still an immature technology, a very high performance level in the first prototype version cannot be assured, but once the basis of the prototype is stated and as the SWS technology matures rapidly these targets should be addressed.

### **4.2.4 Security and performance**

As the system we are developing is a prototype, data security or response time is not an issue of concern. In developing a real GIS system, these aspects are crucial and should be taken into account.

Response time depends on several parameters: what kind of network is used to support the communication, what kind of response is given to the user (plain text, single map, multi-layer maps, etc), and what kind of device the response is represented in (PC, mobile terminal, etc). Another aspect is the accumulated response time of the legacy service/spatial system and the SWS overhead.

The response time is crucial when delivering content to mobile devices, but not so relevant when talking about PCs. Because this prototype is only going to be implemented for personal computers, attention will not be paid to this issue. Response time will be an important issue however for possible future implementations on mobile devices.

Security is crucial if delivering or displaying any personal information. The kind of encryption used also varies depending on what kind of channel is used in the communication (e.g.: current standards suggest a minimum of 128bit encryption across mobile delivery channels)

A balance has to be reached between security and performance. If work on mobile channels results in encryption slowing down data transfer significantly, the graphics and the amount of data delivered should be kept to a minimum (only low density maps or purely textual responses).

### **4.2.5 Different access rights for service consumers**

There are different access rights for service consumers depending on their role. Not everyone accessing the system will have the same rights. Depending on their role (normal users, administrators, editors, etc) they will be able or not to perform certain actions. For example: a municipal intranet website may offer more information than a public client (e.g. an address search or a point of interest database).

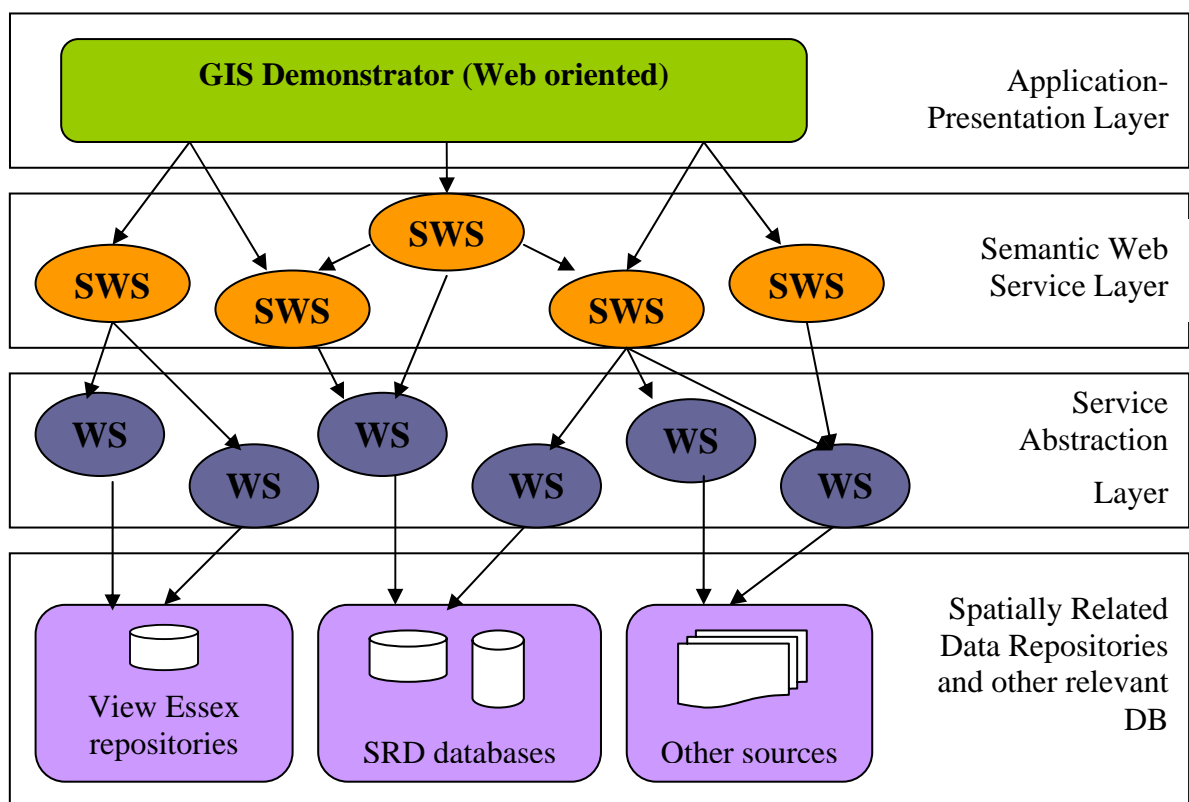
**[NFR 3] Users' roles.** The prototype will be targeted toward the needs of politicians and decision makers in the first stage. There will only be a limited group of users, with the same rights to access data. If the prototype is extended and offered to general public, some of the options might be cut down to preserve access to sensitive data. No one will ever be able under any circumstances to modify the underlying data. SWS only perform queries and operations with the data, not modifying it by any means.

#### 4.2.6 Availability

As the focus is on demonstrating how ODGIS can provide advantages over plain GIS, this prototype will not pay attention to availability requirements. If the system is finally implemented, this issue should be taken into account, as a good eGovernment application needs to be robust and have high availability.

## 5 PROTOTYPE

This section illustrates explains the architecture of the GIS prototype. The picture below outlines the overall architecture, setting the limits of the different layers:



**Figure 10: System prototype**

## 5.1 Architecture

The GIS-prototype is divided into four layers (see Figure 10).

1. The lowest layer is formed by already existing Spatially Related Data (SRD) repositories, already available and in use in ECC and other organizations. This is the information source where data is to be extracted from
2. The second layer is built on top of the SRD systems. It consists of standard Web Services (WS) that access the SRD data repositories, extracting certain relevant data, useful for the application purposes. There should be some services which could perform some basic actions, namely:
  - One or more of these WS have to provide maps covering the area of interest at any moment. These maps will be the basis over which the results of the users' desires will be represented.
  - Some other services should provide coordinates given a certain post code address. This process is called geocoding or address cleaning, which verifies the identity of the address and finds its correct geographic location. This service is crucial in case the retrieved information does not refer to coordinates (so that it is impossible to display on a map)
  - Also some services which perform 'reverse geocoding' might be needed. This process transforms a given position into a description of a feature location, such as a street address
  - Finally, some services to change from one system of coordinates to another one should be developed.

These services are heterogeneous in terms of the data held and the kind of services offered to the user. There can also be several services offering the same service but in different formats. This is the main reason why SWS are needed: to cope with these heterogeneous services.

3. The third layer is the 'core' layer. It is the one that gives semantics to the system. This layer creates SWS on top of the standard WS provided by the Service Abstraction Layer. The second layer assists the SWS layer to manipulate real data in the Data repositories. SWS are to be created with the IRS-III (Internet Reasoning Server), a WSMO [23] compliant server created by the OU. This server imports the GIS domain ontology, and invokes the WS to create SWS that satisfy the user needs (expressed by goals). It is a platform for creating SWS based on WSMO [23] descriptions. This process is totally transparent to the user.
4. The top layer is the one interacting with the user. It displays the results obtained from the execution of the SWS, and displays them in a graphic and web-oriented format. It is a server application that combines the base maps (obtained from the basic WS) with the desired information gathered from other WS. This layer will access IRS-III capabilities through its API<sup>1</sup>. The users can specify information against a background map, by clicking on a certain point or selecting among a list of options. The background map (containing basic elements like roads, rivers, railways, etc) is the first

requirement of any mapping project. The information selected by the user, along with the point or the area of the map chosen, will be given to the IRS-III server as the goal. The goal will unleash the execution of the correspondent WS. The result of the WS will be presented back to the user. The format in which this information will be presented to the user depends on the kind of service executed; it could be as text, overlapping maps or other kind of representation.

## 6 CONCLUSION

In this deliverable we have shown the overall purpose of the system to be created: to enhance GIS with the use of SWS. The prototype to be created will have the advantages of the dynamic and distributed access of the web services and will also benefit from automatic discovery and invocation due to the semantics.

Building SWS on top of the WS will provide the ultimate functionality to the GIS. Data will be self-contained and self-described, which enables machines to discover and invoke it automatically. Therefore, no advice or any help will be needed to integrate data from different sources.

Within SWS everything is automated, SWS are automatically discovered over the Internet. Each organization provides the data to anybody that wants to access it in any part of the world. These data are accessed automatically and as the data comes with metadata, the client knows what the data is about and whether it meets their needs.

The prototype to be created will be based in WSMO [23] and its four elements: WS, Goals, Mediators and Ontologies. Forthcoming deliverables will illustrate in detail WSMO [23] descriptions (of goals, mediators and WS) and the domain Ontologies to be created to support the overall architecture.

The focus of this system is primarily on the needs of politicians and decision-makers, so they can estimate the consequences of their decisions (e.g.: putting an airport runway in a certain location)

The first line of approach is to focus on statistics and forecasting information rather than a 'response' system (for example, identifying an optimal way to travel at a given time etc...), but this might vary as the use case discovers what data sources are available and what relevant data they contain.

---

## REFERENCES

- [1] ECC (13 February 2004). *GIS Business Case Issue 1.1. and Technical Appendixes*
- [2] Simon Bennett, Steve McRobb and Ray Farmer (2002). *Object-Oriented Systems Analysis and Design Using UML*.
- [3] Frederico Fonseca .*GIS\_Ontology.com*. University of Maine, Orono, USA
- [4] T.R. Gruber (1993). *Towards principle for the design of Ontology used for knowledge sharing*. In N. Guarino and R.Poli, editors, *Formal Ontology in Conceptual Analysis and Knowledge Representation*, International Workshop on Ontology. Kluwer Academic.
- [5] Christoph Ruther. *Specification of Terms to share Heterogeneous Information*. Institute for Geoinformatics, University of Munster. Germany.
- [6] Frederico T. Fonseca, Max J. Egenhofer, Peggy Agouris, Gilberto Camara. *Using Ontologies for Integrated Geographic Information Systems*. National Center for Geographic Information and Analysis and Department of Spatial Information Science and Engineering University of Maine, National Institute for Space Research (INPE) Image Processing Division, Brasil
- [7] Heiner Stuckenschmidt. *Semantic Translation of Land-Use classifications: A case study*. Intelligent Systems Group, Center for Computing Technologies, University of Bremen, Germany.
- [8] Liuming Lu, Guojin Zhu, Jiaxun Chen (1882). *An infrastructure for E-government Based on Semantic Web Services*. Dept. of Computer Science, Donghua University, West Yan'an Rd, Shanghai 200051, China.
- [9] <http://www.opengeospatial.org>
- [10] *Internet Engineering Task Force*; <http://www.ietf.org>
- [11] *World Wide Web Consortium*; <http://www.w3c.org>
- [12] Object Management Group. <http://www.omg.org>
- [13] International Organization for Standardization. <http://www.iso.org>
- [14] Geographic Information/Geomatics .ISO/TC 211. <http://www.isotc211.org>
- [15] <http://www.opengeospatial.org/initiatives/?iid=7>
- [16] Gurber, T. A translation Approach to portable Ontology specifications. *Knowledge Acquisition*. Vol 5. 1003. 199-220
- [17] NIWA and Wandermann . LBScult Technical report. <http://www.lbscult.org/>
- [18] <http://www.pbh.gov.br/prodabel/cde/publicacoes/2000/fonseca2000.pdf>
- [19] <http://www.ingentaconnect.com/content/klu/amai/2002/00000036/F0020001/00406341>
- [20] <http://lbd.epfl.ch/e/teaching/SlidesST/1>
- [21] [http://www.spatial.maine.edu/~fred/fonseca\\_acmgis.pdf](http://www.spatial.maine.edu/~fred/fonseca_acmgis.pdf)
- [22] [http://www.dpi.inpe.br/gilberto/papers/semantic\\_granularity.pdf](http://www.dpi.inpe.br/gilberto/papers/semantic_granularity.pdf)
- [23] <http://www.wsmo.org/>
- [24] <http://www.wsmo.org/wsm1/>
- [25] <http://www.wsmo.org/wsmx/>
- [26] <http://wsmo4j.sourceforge.net/>

## Annex 1 - State-of-the-art in Geographic Services

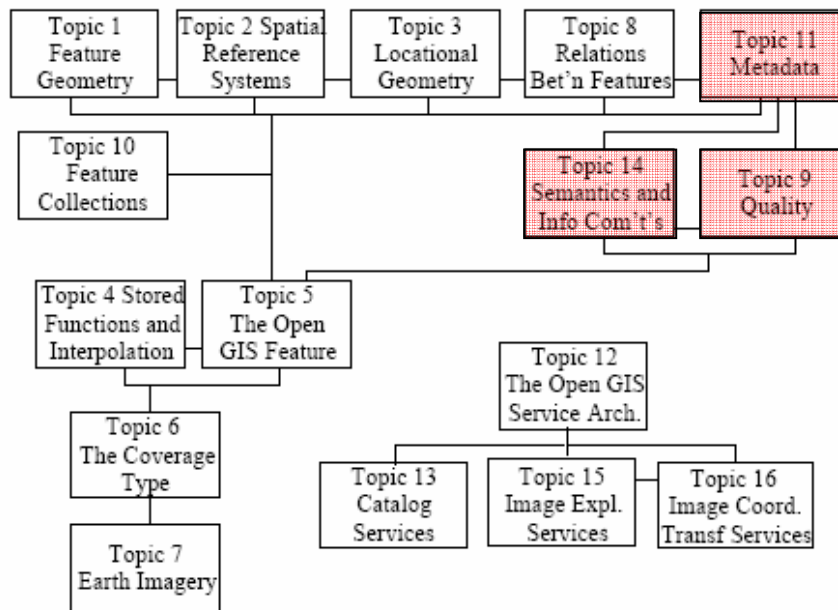
In this chapter we present a short introduction to core technologies and standards in the domain of GIS and location-based services. GIS standards are created by the Open Geospatial Consortium (OGC), but there are some other core technologies in the market. [17]

### **Open Geospatial Consortium**

The Open Geospatial Consortium (OGC) [9] started with the vision of diverse geo-processing systems, which could interoperate over the Internet based on their open interfaces defined in the "Open Geodata Interoperability Specification" (OGIS). Beginning in 1994 with only eight members the OGC has currently more than 250 members from governmental, academic, and private sector organizations. In the following years the OGC developed a set of specifications that together formed a reference architecture for GIS. These specifications were widely adopted by commercial and open source software vendors. As standards for geo-processing cannot exist just on their own, the OGC utilizes standards from other organizations like the IETF [10], W3C [11], OMG [12] or ISO [13]. The most important partnership is an agreement with the ISO TC211 [14] that allows a technical alignment between the partners. Furthermore OGC's specifications can be viewed as implementations of the ISO TC211 standards. On the other hand OGC's specifications will become ISO standards if they meet certain criteria. With other topics like the Geometry Schema, which is the basic definition for geometric objects, there are three standards organizations involved.

The OGC produces two different kinds of specifications: Abstract specifications and Implementation specifications. Abstract specifications define a conceptual model that is sufficient enough to create an implementation specification. The Geographic Information Software products can be built based on this implementation specification. Certain specifications and also standards from other organizations are organized in so-called bookshelves, for example the OpenGIS Abstract Specification Bookshelf. Each of the books in the bookshelf has its own Topic and a descriptive name.

Currently there are 17 books in the Abstract Specification Bookshelf. They cover primarily two technology themes: Sharing Geospatial Information and Providing Geospatial Services. Next figure (Figure 11) was taken from Topic 0: Abstract Specification Overview below shows these topics and their relationships.



**Figure 11: OpenGIS Abstract Specification Bookshelf, Topic 0**

Our initiative can be placed in under the Topics 9, 11 and 14 (which are strongly related). The GIS Ontology to be developed will provide the necessary Metadata to create the semantic relationships among the different GI elements, leading therefore to an enhanced quality product.

In 2001 the OGC has started the **Web Services Initiative (OWS)** [15] to extend and modernize existing standards to enhance their robustness and interoperability. Activities include the adoption of the common architecture for the Web Service Architecture of the W3C including SOAP and WSDL, workflow definitions, increased interoperability for GML and Web Feature Service and enhancements and additions to the Open Location Services.

In DIP we go a step further. We have explored the limitations of Web Services and proposed new standards for SWS, which benefits have been stated all along this document. The ECC-DIP team (e-government case study) is going to apply this new SWS technology to GIS. If successful it might be considered as a future GI standard.

### Other Current Core Technologies

Besides the OGC standards, there are some other core technologies about GIS at the time of writing. Some of these are:

### GPS (Global Positioning System):

A worldwide radio-navigation system of geosynchronous satellites; which continually transmit their position to receiving devices on the earth. This allows computing positions on the Earth with great accuracy by the use of an electronic receiver.

The satellites are used as reference points. There are currently 25 satellites in 6 orbital planes at 20,000km altitude with 12hour periods. They can provide highly precise position, velocity and time information to users anywhere on Earth.

It was developed by the US. Department of Defence, but it's now widely used in marine and terrestrial navigation, personal tracking and automatic vehicle location technologies. Due to its great precision it is used to support cadastral surveying.

### Oracle Spatial:

Oracle supports geographic data types in its core product and this option is becoming widely adopted for supporting corporate spatial data warehouses. The standard option is known as *Locator* and this supports most basic geospatial data storage needs.

The full functionality comes as a component of Oracle Enterprise Edition that provides advanced spatial features to support high-end GIS and LBS solutions.

This is a commercial solution which has been widely adopted by many GIS and LBS vendors. It has also been deployed by telecommunications, utilities, and government organizations worldwide.

It provides a robust foundation for complex GIS applications which require more spatial analysis and processing in Oracle Database. It includes spatial functions (including area, buffer, centroid calculations), advanced coordinate systems support, linear referencing system, and aggregate functions

### Location Based Services (LBS):

It refers to a kind of services which can be offered locally based by the access to the respective position coordinates, rather than being a kind of technology.

LBS comprise the functionality needed for offering services and content to a mobile user. This includes services like: locating a device in a mobile network, coordinate transformations, communicating with a mobile device, interfacing with mobile network operators, integration of third party services (e.g. map generation, address clearing).

## Annex 2 - Questions' template for the interviews

- **General Questions (Real-world problems)**

### **General purpose of your system:**

1. Could you please briefly describe your use of geographic information? Describe the nature of the contents of your GIS and their level of detail.
2. What problems are you trying to solve (that require use of data)?

### **Kind of data:**

3. What data do you use?
4. Does this data belong to the ECC or could be accessible to us? Does it have any data-sharing restrictions (legacy issues, privacy data...)?

### **Kind of operations with data:**

5. What kind of operations do you perform with the data? (What kind of questions do you ask to the system?) (What kind of functionalities does your system offer?)
6. How do you access it? How do the users access the data? Through a specific program? Through a web browser?

### **Authorised users:**

7. How many people can access your data?
  1. What are their roles? Manager, supervisors, operatives...
  2. Can all they perform the same operations (do they have different permissions)? Can any of them modify the data?
  3. What are the expected inputs from the user?
  4. And the outputs? (How do you display the results? Plain text? Overlapping maps?)
8. Who administrates the data and in what intervals?
9. What problems do you experience in pulling it together for Decision-making purposes?
10. What additional functionalities would you like to add to your system?

### **Where is data kept?**

11. Where is it stored in? (Databases, text, maps, pictures ...)

12. Is it stored in one or more than one data repositories?  
 If in more than one:
1. How do you merge this data?
  2. Do some organizations (or systems you are trying to merge) manage data differently than others? Which ones? How do you cope with this?
13. Do you only work with data stored in Geographically Related Data repositories or also with any other data sources (web pages, free texts, and any other kind of Databases (healthy DB, social data, economic data, and historical data)...)?

### **Merging data:**

14. Have you merged your data with other data sources?
1. If so: How do you consolidate/reconcile data between the systems (when merging them)?
  2. If not: If not: would you like to?

Would you consider it worthy to share information between ECC and other bodies? (Databases from other departments or companies) e.g. District Councils, Emergency Services, Central Government etc. Are there any sources you'd consider useful to merge with your system? (Free texts, web pages, data in non-electronic formats...etc)

### **Heterogeneity problems:**

Have you experienced any heterogeneity problems (data inconsistency)? Like the following:

15. When the numbering of a street changes: Do you store somewhere the previous numeration so users can still look up by the previous number?
16. Do you cope with popular names? Some places addressed by people by its "popular" name instead by its official name (when an ancient name or a popular name is more known and used than the official one). Do you store this information?
17. Implicit meaning of data: What does "main road" mean to you? (Width, number of cars, longitude...?) **Is the meaning of the fields stated somewhere or it is implicit knowledge?**
18. How do you address roads? By name or by number? What would happen if the user wants to look for a road using the other classification?
19. Have you already solved these problems? How?

**Real time data:**

20. Do you work with real-time data? (Data that is constantly changing through time): weather, emergency events: floods, fire, traffic...?
21. What statutory responsibilities do you have relevant to geographic information... what standards do you have to comply with for your outputs?

- **More Technical Questions**

22. What data attributes are important? Can we have a sketch of the DB structure? Do you have any description of the fields (their names and data-types)? Are there any mandatory fields?
23. Give a short overview of applications based on your GIS system.
24. Based on your experience: where lie the problems when having to integrate or cooperate with other (GIS or non-spatial) applications? And what type of systems have been integrated so far?

Questions that appear in some manner already in the list of questions above:

25. Is the system coupled with other systems (spatial or non-spatial) e.g.:
  - address search/clearing
  - databases from other departments or companies
  - geo-coding/referencing
26. What has to be done when adding a new database to the system (in terms of relating the data records to the GIS system)?

- **Very Technical Questions**

27. Do you use standardised tagging/annotation schemes or thesauri?
28. What GIS-Software do you use? Do you currently use an application-specific bespoke GIS system, an off-the-shelf standard GIS tool like ArcGIS or MapInfo, or do you just use basic office tools and manually created paper maps?
29. Can the information be accessed through Web Services?
30. Is the information available only for computers or also for any other electronic devices (e.g.: PDA)?
31. Do you follow any GIS standards? XML maps? Do you follow the convention of OpenGis?
32. What services are supported by your GIS (WMS, WFS, ...)
  1. are they implemented completely
  2. how can they be accessed (HTTP, SOAP-WS, ...)
  3. in what way can the client customize the results
  4. Are any security or financial procedures established (e.g. SSL, billing etc.)

Annex 3 - Metadata schema

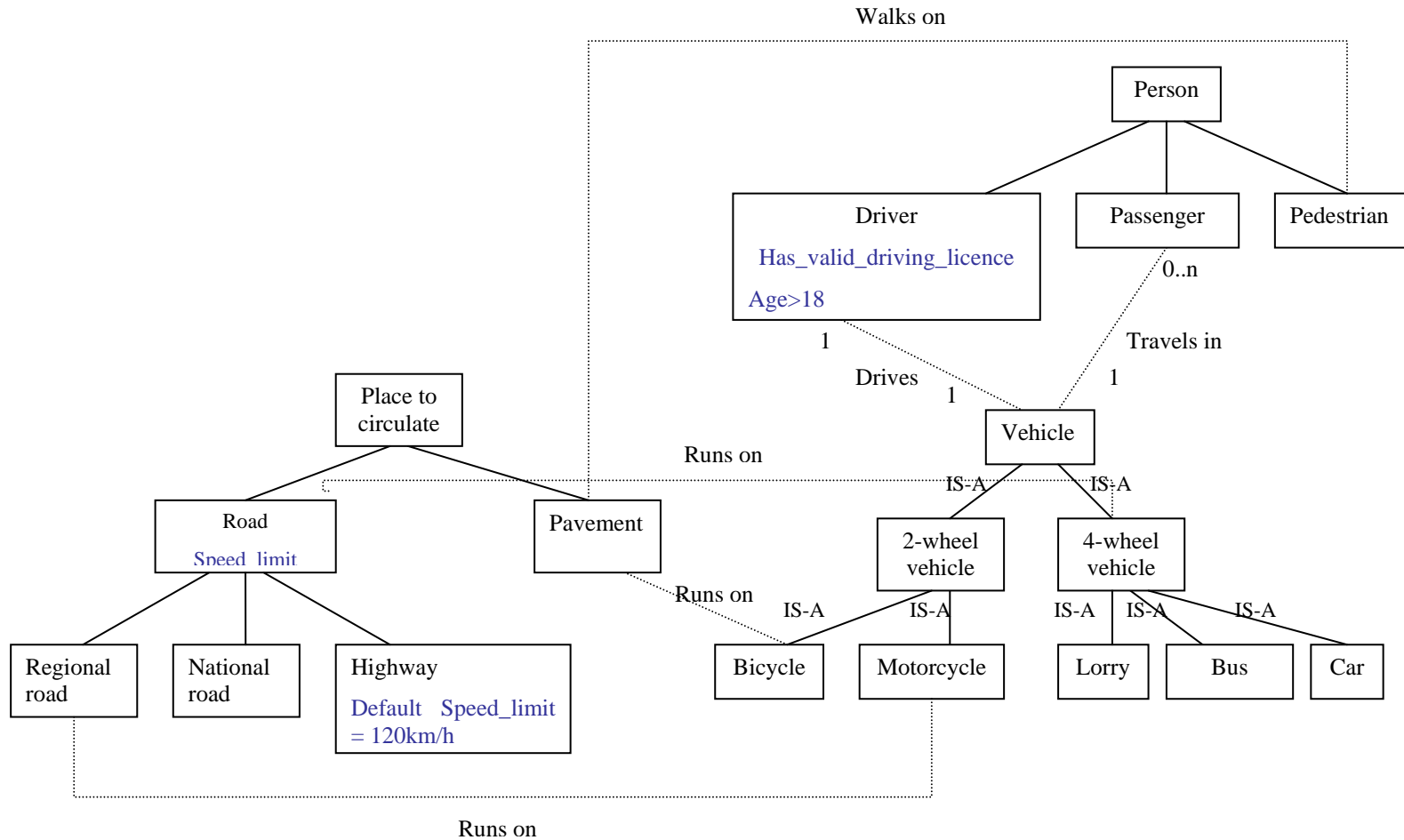


Figure 12: Ontology schema example