eChallenges e-2011 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2011
ISBN: 978-1-905824-27-4

Using Cloud Computing to Increase the Usability of Distributed Web GIS Systems

Stephan GRUIJTERS[1], Steven FRUIJTIER[2]
[1]*TNO Geological survey of the Netherlands, Princetonlaan 6, Utrecht, 3508 TA, The Netherlands*
Tel: +31 88 866 42 56, Fax: +31 88 866 44 75, Email: stephan.gruijters@tno.nl
[2]*Geodan, President Kennedylaan 1, Amsterdam, 1079 MB, The Netherlands*
Tel: + 3120 5711 311, Fax: +31 20 5711 333, Email: steven.fruijtier@geodan.nl

Abstract: Europe's industry is depending heavily on raw materials and energy, mostly imported from countries outside Europe. To secure the future supply of these commodities a distributed web portal will provide online access to harmonised, INSPIRE compliant data on energy and mineral resources, using web services based on open source standards. An internet questionnaire provided user requirements. Together with an inventory of available data a set of key economic attributes is derived. The services are implemented in the cloud, assuring performance and advanced portal functionality, resulting in a Google like user experience.

1. Introduction

Access to and affordability of mineral raw materials is crucial for the sound functioning of the EU's economy. Sectors such as construction, chemicals, automotive, aerospace, machinery and equipment sectors which provide a total value added of € 1 324 billion and employment for some 30 million people all depend on access to raw materials [1]. A large part of these raw materials are produced by countries outside Europe, making access to these commodities vulnerable to the economic and political developments in these countries. EU authorities currently compile their long-term policies regarding the need for oil, gas and minerals, including estimates of the required import, from national reports contributed by the member countries. These reports contain only generalized information regarding reserves and production forecasts for a country as a whole and do not allow a fast response to crisis situations and significantly reduce the accuracy of the long-term planning of the geo-energy supply of Europe. Furthermore the way the data is available; users can not query the data interactively, or combine it with their own spatial datasets for more complex analysis.

Europe has to secure its energy and non-energetic minerals supply for the future. To address this issue the EuroGeoSource (EGS) project is building a multilingual web GIS system that will allow users to identify access, use and reuse aggregated geographical information on geo energy and mineral resources, provided by geological surveys from at least ten countries in Europe. The system will enable OGC-compliant services [2] for the registration of INSPIRE [3] compliant data sets from different countries, the visualization and overlay of the information layers obtained from distributed sources, spatial analysis, and so on. At the portal 11 Geological Surveys will provide their data, and other Surveys are encouraged to share their data to maximize the coverage of Europe.

2. Objectives

This paper will discus the results of the first year, including an inventory of user needs and the derived use cases for the portal, an overview of data available at the surveys involved and the data model composed from this. Furthermore, apart from the users appreciation of the general design of the portal and it's functionality, the technology used within the web portal, in combination with the lessons learned from being a pilot project for INSPIRE are discussed.

3. Methodology

In order to get a good grip on the user needs an on-line internet questionnaire was designed. The questions related to what kind of data users (want to) use, how they (want to) use these data, how they currently acquire their data, with what kind of information they would like to combine it and what kind of information and functionality they would like to have present at the portal. The questionnaire contained questions with answers (1) constrained to a checklist, (2) gathered through free text entries and (3) constrained to a scale of appreciation ranging from 1 (not agree / important) to 5 (agree / important). The questionnaire was sent out to 1041 (potential) users and the results from 187 respondents were analysed statistically, creating a basis for use cases for the design of the portal. The inventory of the data available at the 11 Geological surveys involved was performed by a questionnaire (answers in free text and by checklists) and oral interviews with the data providers. A set of key economic parameters was derived by combining the user requirements with the data available, organising them into a comprehensive data model. To ensure on the spot feedback from the user community data providers, representatives from the mineral and energy industry, international policy makers and INSPIRE experts were invited to a workshop where the rationale behind the portal was explained and the first version of the portal itself was presented and discussed.

4. Technology Description

4.1 Services

The harmonised data model is implemented at all geological survey participating in EGS and is based on and tuned with the INSPIRE themes geology [4], mineral resources [5], energy resources [6] and administrative units [7]. In order to distribute and portray the data three type of services are used. One service for filtering and requesting the data itself (Web Feature Service: WFS), one for discovering the data (Catalog Services for the Web: CSW) and one for portraying the information as a geographical map (Web Mapping service: WMS). Furthermore translation services will be implemented to translate the geological and user interface terms to different languages, as well as processing services to aggregate information. In order to ensure interoperability between the geological surveys and the different system components, all interfaces are based on open standards from the OpenGeospatial Consortium and the INSPIRE specifications, but additionally also on standards such as GeoJSON [8] and Rest [9] resulting in a better integration of the system components in other information systems. E.g. clients for the Microsoft Surface in order to accommodate collaboration and decision making will be developed.

4.2 Cloud Computing

Although the data originates from services and databases installed at the different participating Geological Surveys, cloud computing is used to fulfil some basis requirements, both non-functional and functional. Typical non-functional requirements are the performance, availability and scalability. A user of the system may search for occurrences of commodities throughout Europe, e.g. to explore the possibilities to mine rare earth materials. In case all information is available on distributed servers, such a query will have to be executed at every geological survey, resulting in a high risk of low performance. Therefore the data is stored centralised to act as an optimised search index. It also reduces the risk of having actually inaccurate results if local services are down or unreachable. Furthermore public data available elsewhere could be added on a lower scale (e.g. per country instead of per occurrence) thus stimulating the non-participating countries to join and to give more complete results of the system in the meantime. Functional requirements flourish by centralised data storage. For example in order to be able to return numbers related to hydrocarbon and mining resources per country and per commodity. Otherwise the data retrieved from the distributed services will have to be grouped and summed in the client again with the possibility to have inaccurate data returned to the user due to services being down.

The system also uses cloud computing to compute so called tiles of the WMS services. A typical WMS setup is only able to support a very limited number of concurrent users and requests per second as it creates a map per request. Therefore using a WMS often results in poor usability if many users are accessing the system or if the system is over-requesting the individual WMS. The data in the database is actually not changed very often (a few times each year). Therefore it is possible to create and cache tiles of WMS maps (similar to Google Maps and Bing Maps technology). In EGS the creation and caching of tiles is implemented in the cloud (Amazon Elastic Web) giving flexibility to the resources needed to pre-create the tiles and have a dynamic storage size for the tiles. Also the system is automatically scalable if the number of users and requests rises (expected or unexpectedly).

Another reason to (pre-)create and store tiles in the cloud is because other existing map services are used. In the case of OneGeology for example, the WMS services support only the geographical coordinate system ERTS89. While this coordinate system is very useful for exchanging information, it is a very poor coordinate system for portraying geographical maps. It leads to significant distortion the further North the data is visualised. In the EGS portal Spherical Mercator is used, a projected coordinates system also used by e.g. Google. This leads to much better looking maps. The transformation from maps based on geographical coordinates to tiles based on projected coordinates, is done using cloud computing, as well as the subsequent storing (caching) of the tiles.

4.3 Functionality

Typically a web portal using spatial information is only visualising the information, either graphically in a map or showing the attribute information per geographical object in a table or pop-up screen. However, contrary to other systems in the case of EGS the emphasis lays not only on making data available, but also on the questions which must be answered by the system and the data. These have been elaborated in use cases and prototypes which have been discussed again with the (potential) users. This has lead to a user interface where a user can easily "ask" a question that gets answered in different ways suiting his needs. For example a user can explore the occurrences of Lithium within a certain deposit type, like Pegmatite, without having to define a query manually, but intuitively and yet with extensive possibilities to define the question. If someone searches for example for a certain commodity, the information is aggregated and presented in different ways automatically using maps and diagrams with aggregated information. The diagrams can quickly show where (in what country) the most potentially rich areas are, what the most common deposit types are, etc. The diagrams are also interacting so a user could quickly zoom to Portugal by clicking in the diagram if he sees there are a lot of potential occurrences there.

5 Results

5.1 Current Data Management

At present the content and structure of information on mineral and energy resources differs substantially in each country [10]. Sometimes the information is stored in one national database managed by one institute (Hungary, the Netherlands and Poland). In other countries more then one national institute is responsible for energy and mineral resource data (Bulgaria and Estonia). Finally the information was found to be distributed over several regional databases (Belgium). This makes it difficult for users to know if they have found all available data, if the information is up to date and understand the quality and reliability of information. Using the available information is further complicated because data must be acquired in different (digital) formats and at different levels of processing. Table 1 shows where data related to minerals and energy resources can be found.

Table 1: Active Websites with Information on Geology, Mineral and Energy Resources in Europe

Website	Purpose
Belgium	
www.naturalsciences.be/geology https://dov.vlaanderen.be/dovweb/html/engels.html http://environnement.wallonie.be http://environnement.wallonie.be/cartosig/cartegeologique	Web site of the Geological Survey of Belgium Database for the subsoil and soil of Flanders Regional environmental geoportal of Wallonia Geological map of Wallonia
Bulgaria	
www.mi.government.bg www.moew.government.bg www.mrrb.government.bg	Web site of the MEET Web site of the MOEW Web site of the MRDP
Estonia	
www.egk.ee http://geoportaal.maaamet.ee/eng www.energia.ee/en www.envir.ee www.gi.ee	Web site of the Geological Survey of Estonia Geoportal of the Estonian Land Board Web site of the Estonian Oil and Shale Company Web site of the ENVIR Web site of the TTÜ GI
Hungary	
www.mbfh.hu www.mafi.hu	Web site of the Hungarian Min. and Geol. Office Web site of the Geological Institute of Hungary
Poland	
www.pgi.gov.pl http://geoportal.pgi.gov.pl/portal/page/portal/PIGMainExtranet http://geoportal.pgi.gov.pl/portal/page/portal/MIDASGIS/start	Web site of the Polish Geological Institute Geoportal of the Central Geological Database Geoportal of MIDAS
Portugal	
www.lneg.pt http://e-geo.ineti.pt http://www.dgge.pt	Web site of the LNEG Geoportal of geological databases Web site of the DGEG
Romania	
www.namr.ro www.igr.ro	Web site of the NAMR Web site of the Geological Institute of Romania
Slovenia	
www.geo-zs.si http://akvamarin.geo-zs.si/ms/	Web site of the Geological Survey of Slovenia Web application for mineral resources
The Netherlands	
www.nlog.nl www.dinoloket.nl www.thermogis.nl	The NL Oil and Gas Portal DINO Portal Geothermal Energy Portal

5.2 User Needs

It appeared [11] that the availability of GIS software is not common for a large user group. Having basic GIS functionality available at the portal (e.g. search by keyword and by geographic location) is therefore highly appreciated, with an average score of 4.5, whereas more advanced queries on the available data are also important (average score of 3.7). Concerning the possibilities for viewing data, there is a slight preference for the presentation of data on maps (4.6) and in tables (4.4), rather than as graphs (4.1) or by viewing the original document (4.2). Downloading data is considered an essential functionality (average score 4.7). Users want information on metallic, non metallic and energy resources (see fig. 1), and want to be able to combine it with other relevant spatial on geology, ecology, basic geography, land-use and economic data.

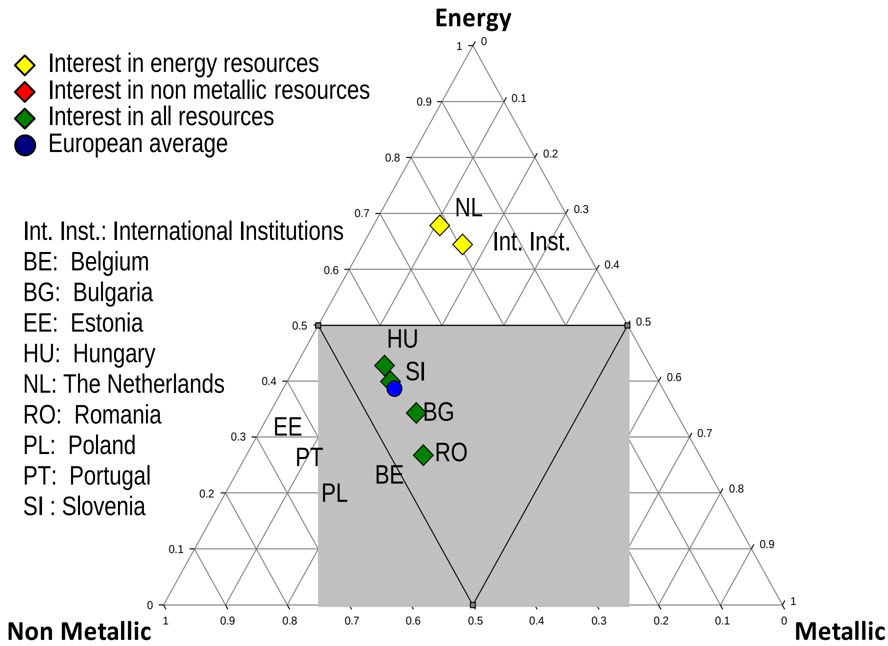

Figure 1: Questionnaire Results: Importance of Data on Energy/Metallic/non metallic Resources Per Country

The data provided by the EuroGeoSource portal will be used in the context of industrial planning, economic issues (analysis and forecasting) and the mineral policy area. Experts from OECD, DG Enterprise and Industry, EESC, JRC, Euromines, OGP, University of Luleå and the Mining Authority from Portugal confirmed this outcome [12], stressing the importance of collaboration between the Geological Surveys of Europe and sharing best practices to increase EU's knowledge base.

5.3 EGS Key Attributes

The set of key attributes [13] is grouped in general data of the site, data of location, administrative data, economic data and additional data (see table 2). Attributes specific to mineral resources are labelled (M) to energy resources (E) and to both (M&E). Not all attributes are available and or accessible for each country. The data are mapped to INSPIRE data specifications [3, 4, 5, 6], leaving twenty-one attributes that could not be mapped. The mapping showed that the connection between the different data themes and the rationale within each theme in INSPIRE is not yet optimal, resulting in multiple references our redundancy when entering data. These insights will be reported to the INSPIRE thematic working groups in detail to improve the final versions (3.0) of the data specifications.

Table 2: Key attributes on EGS portal for mineral resources (M), Energy resources (E) or both (M&E). The shaded attributes are additional to the INSPIRE data model.

GENERAL DATA OF SITE	DATA OF LOCATION	ADMINISTRATIVE DATA
INSPIRE ID of site (M&E)	Coordinates: longitude, latitude (M&E)	licence ID (M&E)
local ID (M&E)	depth below surface (M&E)	type of licence (M&E)
Name of site (M&E)	water depth (M&E)	Name of licencee / operator (M&E)
Name of site (M&E)	geographical location (M&E)	Duration of licence (M&E)
Type of resource (M&E)	Country name (M&E)	Areal extent of licence (M&E)
Year of discovery (M&E)		
status of site (M&E)		
References (M&E)		
Remarks (M&E)		
ECONOMIC DATA	**ADDITIONAL DATA**	**ADDITIONAL DATA**
Classification (M&E)	Geological characteristics regional / of field (M&E)	Main type of field (E)
in situ ore / substance reserves (M&E)	Age of host rock / Reservoir rock age (M&E)	Status (E)
Production (M)	Host Rock type / Reservoir rock type (M&E)	Nr of oil producing wells (E)
Period of Production (M)	mineral deposit type (M)	Nr of gas producing wells (E)
Dimension of the deposit (M)	primary comodities (M)	Nr of gas injecting wells (E)
mining method (M)	secondary comodities (M)	Nr of oil/gas producing wells (E)
Oil Initially in Place (E)	main ore minerals / substance (M)	Nr of water injecting wells (E)
Gas Initially in Place (E)	secondary ore minerals / substance (M)	Nr of water/gas injecting wells (E)
Cumulative oil production (E)	(M)	Nr of CO2 injecting wells (E)
Cumulative gas production (E)	hydrothermal alteration (M)	Nr of producing/injecting wells (E)
Cumulative water production (E)	morphology of the deposit (M)	Areal extent of field delimitation (E)
Cumulative gas injection (E)	regional deposit structure (M)	Reservoir depth (E)
Cumulative water injection (E)	dating method of mineralisation (M)	Production strategy (E)
Remaining Oil reserves (E)	age of mineralisation (M)	Installations (E)
Remaning Gas reserves (E)		
Year of reporting (E)		

5.4 EGS Architecture

In order to accommodate the desired functionality and performance of the EGS portal the architecture of the EGS system is divided into three layers (Figure 2):

- Data / service provider layer, containing components responsible for delivering data to the central service layer.
- Central EGS service layer, providing the services that are required by the consumer layer and for caching the data delivered by the provider layer.
- Service consumer layer, containing components that consume EGS services (the EGS portal and external client applications).

The harmonized EGS data will be harvested by a harvesting service at regular intervals, during 'low activity' hours and stored in the feature cache data store. Furthermore, the data providers can manage the harvesting. Also non-harmonized data, provided through WMS services, can be used to generate tile caches at regular intervals. The user can see when the latest update of the data was performed. In this way the user gets access to reliable, accurate, up to date information from the data providers, without having to wait for all the services to load at the portal.

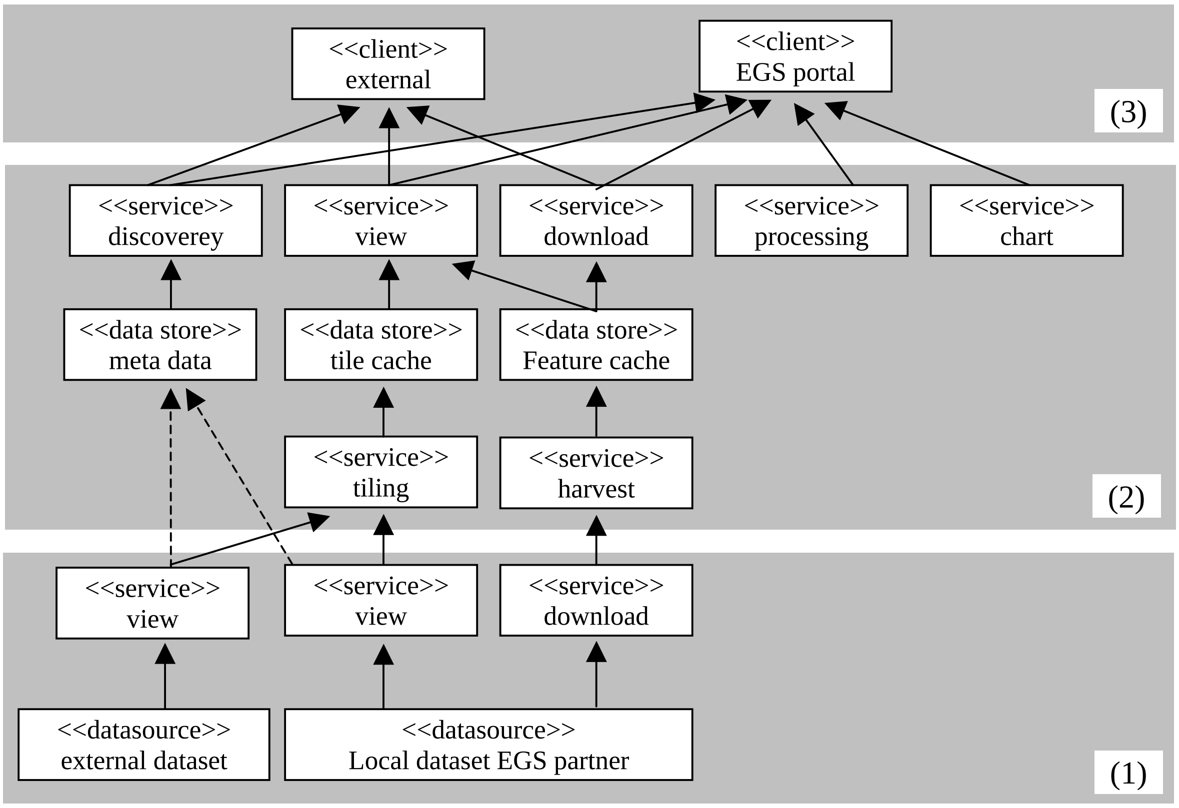

Figure 2: The EGS architecture with (1) data/service provider, (2) central EGS and (3) service consumer.

6 Conclusions

The taken approach of extensively discussing the functionality with the end users and using techniques such as cloud computing and (tile) caching leads to an innovative but most importantly well performing, user-friendly and useful system available in several languages with a high degree of flexibility.

The design of the EGS portal shows how distributed web services are indeed essential to search, provide access, use and re-use INSPIRE compliant data. But to keep the performance of the portal acceptable, cloud computing is necessary.

The development of additional functionality, providing the user with aggregated information form the EGS data model, is an important asset of the portal: in stead of only providing harmonised data per country it integrates the data on the European level, making it possible to address European supply and policy issues.

At this moment the number of connected geological surveys is limited to a few test surveys within the consortium. Once a critical number of geological surveys is connected to the system, it will be a indispensible tool for companies and organisations on regional, national and European level working on mining and energy resources.

References

[1] COMMISSION OF THE EUROPEAN COMMUNITIES, The raw materials initiative — meeting our critical needs for growth and jobs in Europe, COM(2008) 699 final, Brussels, 2008
[2] Open Geospatial Consortium Inc., OGC reference model, OGC 08-062r4 version 2, 2008
[3] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union, 2007, pp. L108/1 – L108/14.

[4] INSPIRE Thematic Working Group Geology, D2.8.II.4 INSPIRE Data Specification on Geology – Draft Guidelines, version 2.0, 2011
[5] INSPIRE Thematic Working Group Energy Resources, D2.8.II/III.20 INSPIRE Data Specification on Energy Resources – Draft Guidelines, version 2.0, 2011
[6] INSPIRE Thematic Working Group Mineral Resources, D2.8.II/III.21 INSPIRE Data Specification on Mineral Resources – Draft Guidelines, version 2.0, 2011
[7] INSPIRE Thematic Working Group Administrative units, D2.8.I.4 INSPIRE Data Specification on Administrative units – Guidelines, version 3.0.1, 2010
[8] H. Butler et al., The GeoJSON Format Specification, version 1.0, 2008
[9] R.T. Fielding, Architectural Styles and the Design of Network-based Software Architectures, University of California, Irvine, 2000
[10] P. Scharek, T. Tullner, Political and organisational aspects of geo-energy and mineral resources data management in the participating countries, EuroGeoSource report 3.1, Budapest, 2011
[11] N. Maricq, K. Piessens, Requirements for geo-energy and mineral resources data delivery at European and national users' levels, EuroGeoSource report 2.1, Brussels, 2011
[12] First international EuroGeoSource workshop, www.eurogeosource.eu, Budapest, 2011
[13] J. Šinigoj, A set of interoperable key attributes for the spatial objects, EuroGeoSource report 4.1, Ljubljana, 2011
[14] S. Fruijtier et al., draft architectural design of the EuroGeoSource system, EuroGeoSource report 6.1, Amsterdam, 2011
[15] F. Waardenburg et al., services for accessing digital spatial information of geo-energy and mineral resources, EuroGeoSource report 5.1, Utrecht, 2011, in prep.