

EuroGeoNames (EGN) – Implementing an INSPIRE service

Pier-Giorgio Zaccheddu^a, David Overton^b, Saulius Urbanas^c

^a Federal Agency for Cartography and Geodesy (BKG), Germany

^b dbyhundred ACIM, United Kingdom

^c EuroGeoNames Project Manager (EuroGeographics), Belgium

Keywords: geographical names; exonym; endonym; web feature service; cloud architecture

Executive summary

The EuroGeoNames (EGN) project funded by the eContentplus programme of the European Commission (EC) started on 1 September 2006 and the funded period lasted until 28 February 2009 with a project budget of €1.8 million. During the funded period a European geographical names infrastructure was established by connecting existing national official data sources of the participating National Mapping and Cadastral Agencies (NMCAs). In this way the geographical names data is updated in a consistent way and maintained at the source level by the responsible organizations.

Recognized as a very successful project by the European Commission and the project partners, EGN then progressed to become the first INSPIRE¹-compliant service in Europe during the implementation phase (to 2012). EGN now provides excellent opportunities for EuroGeographics and its members to meet future requirements in their role as reference information providers for national and European spatial data infrastructures.

Between 2009 and the end of 2011, the overall management of the EuroGeoNames infrastructure was conducted by EuroGeographics together with the German Federal Agency for Cartography and Geodesy (BKG).

¹ In Europe a major recent development has been the entering in force of the INSPIRE Directive in May 2007, establishing an infrastructure for spatial information in Europe to support Community environmental policies, and policies or activities which may have an impact on the environment. INSPIRE is based on the infrastructures for spatial information established and operated by the 27 Member States of the European Union. The Directive addresses 34 spatial data themes needed for environmental applications, with key components specified through technical implementing rules. This makes INSPIRE a unique example of a legislative “regional” approach.

Implementation continues in 2012 as EGN moves closer to its aim to achieve at least EU27 coverage as a part of an initiative to make EuroGeoNames self-sustaining.

A new architecture and management approach for EGN was explored partly within the ESDIN project (a European Spatial Data Infrastructure Network) in collaboration with the EGN Coordination Committee. In this project best practice has been sought for reaching INSPIRE compliance and creating harmonized pan-European location data and services. As Geographical Names was one of the five themes delivered by ESDIN in 2011, there were a number of insights and improvements which are now included in the EuroGeoNames implementation.

Background

The increasing use of public domain geographical information, especially geographical names data, raises an interesting question given the pedigree of such data: who decides where (for the general population) European places are and how they are spelled?

Owing to the multilingualism and cultural richness of Europe, its people use different spellings and languages when talking about the same location and even within a country more than one spelling may be used. This is a European situation that should be considered a matter of pride and not an obstacle. Thus, the full richness, completeness and high quality of European data seem only to be guaranteed if the data providers are the individual European countries – creating and maintaining the source data themselves.

According to existing popular web services, using public domain data sources, the answer would appear to be that by default, and in lieu of a European alternative, it is not the respective National Mapping and Cadastral Agency (NMCA) – or another national institution – that decides where, for example, Bruxelles/Brussels is located and how it is spelled.

Objectives and achievements

Within the EU-funded period (1 September 2006 – 28 February 2009) the EuroGeoNames (EGN) Consortium implemented a web (gazetteer) service infrastructure for providing official geographical names data in Europe, in conjunction with about 20 European NMCAs, to help

one to find the official spelling of a name, together with its spelling in other languages, its geographical location, its pronunciation, etc.

Figure 1: Participating and associated National Mapping and Cadastral Agencies (NMCAs)

Comment [o1]: Place image "2.4_Pier-Giorgio_Zaccheddu_&_David_Overton_1.jpg" below this caption

The official geographical names data kept decentrally in the EU countries is linked to and searchable via so-called 'Web Feature Services – WFS'. A so-called 'EGN Central (WFS) Service' accesses the distributed 'EGN Local (WFS) Services' at each data provider (NMCA) to query the EGN data network and return standardized result sets (in XML) to the inquirer. The data maintenance and updating process remains the responsibility of the countries that have collected and maintained them. An exonyms and other variant names database (EVN-DB) comprises important names used in a specific language for a geographical feature situated outside the area where that language is spoken, and differing in its form from the name used in an official or well-established language of that area where the geographical feature is located. These names are not part of the databases of the participating NMCAs of the EGN project. The EVN-DB is a supplementary database to the EGN Central Service and each (standardized) exonym and other variant name is linked unambiguously with the appropriate official endonym(s) provided by the NMCAs. An online-editing service serves to maintain the EVN-DB in future. Single requests (currently limited to 50 requests per day for anonymous users) for geographical names by using the EGN Central Service are free of charge.

Figure 2: EuroGeoNames services' architecture

Comment [o2]: Place image "2.4_Pier-Giorgio_Zaccheddu_&_David_Overton_2.jpg" below this caption

Generally, the EGN Consortium aimed at being as compliant as possible with the findings of the INSPIRE² initiative. The major strength is that the EGN gazetteer model is compliant with

2 The Directive 2007/2/EC of the European Parliament and of the Council adopted on 14 March 2007 aims at establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) for environmental policies, or policies and activities that have an impact on the environment. INSPIRE will be based on the infrastructures for spatial information that are created and maintained by the Member States. To ensure that the spatial data infrastructures of the Member States are compatible and usable in a Community and trans-boundary context, the Directive requires that common Implementing Rules (IR) are adopted in a number of specific areas, amongst others, for the topic geographical names.

the INSPIRE gazetteer model (which is based on the ISO 19112 standard). As for the final version of the INSPIRE data specification for geographical names, it can be stated that the output of the EGN Central Service can be mapped to the INSPIRE data specification and that EGN currently conforms to all content mandatorily required by INSPIRE.

EGN is targeted primarily at value-added resellers (VARs) and service providers to develop specific applications for their customers and deploy value-added Geographic Information System (GIS)³ products by using the EGN Central Service. The end user has access to this information either through the gazetteer service interface or through the applications (end-user interfaces). Two applications have been developed within the funded period.

The first one is the so-called EGN Reference Application,⁴ which enables searching for geographical names in all official European languages, including officially recognized minority languages, showing the full functionality of the EGN infrastructure. The second one is the EGN ArcGIS extension developed by ESRI. This extension enables ArcMap to perform name searches based on several query criteria and to analyze, visualize and save the results in a standard GIS software environment. The EGN ArcGIS extension is available as a free download through ESRI ArcScripts.⁵ Other applications, developed by third parties, will be encouraged.

Status of the EuroGeoNames infrastructure

Since development of this infrastructure the main task has been to ensure the service can be sustained in the medium to long term. Boosting the number of connections to achieve a “critical mass of content”⁶ continues as we progress toward full EU27 coverage.

Figure 3: Overview of connected National Mapping and Cadastral Agencies and plans for the EU27

³ A geographic information system is a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data. The acronym is GIS, A GIS facilitates the merging of cartography, statistical analysis, and database technology.

⁴ <http://www.eurogeonames.com/RefAppl/> (accessed 30 March 2012).

⁵ <http://arcscripts.esri.com/> (accessed 30 March 2012).

⁶ As stipulated by the eContent funding.

The quantity of geographical names data in the respective countries is variable, and depends on the general shape of the country, the map scales used as the basis for data acquisition, as well as on the level of information detail required in the countries themselves. It varies from 1:5,000 (e.g. Cyprus) and 1:10,000 (e.g. The Netherlands) to 1:200,000 (e.g. Germany). In the initiation phase of EGN, a remarkable amount of time was required for analyzing the data to be aggregated during the project. A gazetteer service allows a client to search and retrieve elements of a geo-referenced vocabulary or database of well-known geographical names or named places in general. Each named place is associated with one or several geographical names, i.e. proper nouns applied to a spatial object. The different geographical names of one given spatial object may be for example the names in different languages or in different forms (e.g. complete and short forms of country and administrative unit names). The scale information is not relevant. It is of more importance that, for example, all geographical names of 'administrative units' are available. The minimum content of attributes for a gazetteer service is: geometry, describing the footprint or a reference point of the named place as well as one or several type(s), characterising the kind of entity represented by the named place (e.g. 'administrative unit').

Currently, the connected German service provides about 90,000 entries, the Finnish one about 808,000 entries. Thus, assuming that about 20 countries mentioned above will be connected to EGN in 2012/2013, the quantity of geographical names may lie in between 2.8 million and 3.5 million entries (endonyms). In addition, about 8,000 exonyms and other variant names will be linked to the national databases.

Target user and their needs

Within the EGN project, the market in terms of potential applications and related business actors and stakeholders, from both public and private sectors, was analyzed, as well as user and business requirements based on the outcome of the market analyses described.

As for the target user groups, the following ones can be distinguished:

- Emergency services; health and safety
- Cross-border market analysis and asset management
- Border crossing routing; transport and delivery service networks
- Hotel reservation services, tourism

- Private-sector map and atlas producers
- Educational establishments, libraries
- Mass media (newspapers, TV)
- Location based services (LBS)

Generally, the JRC INSPIRE geoportal is potentially a user for EGN, as well as agencies of the European Commission, such as Eurostat, and also commercial actors like Google, Yahoo, Microsoft, Tele Atlas, Nokia.

The so-called 'EGN Group of Interest', comprising 26 organizations (standardization bodies, private GI companies, cartographic publishing houses, GI interest groups, etc.) potentially interested in using the EGN infrastructure and services too, was invited to join, comment and shape the EGN project activities from the beginning in 2006 and to provide 'customer feedback' by participating in workshops.

The unique selling points (USP) and qualities of the EGN infrastructure and services for Europe are that:

- The names data provided are from a primary source.
- It is continuously updated.
- It is more detailed than other names data.
- It is closer to the experts that collect the names.
- There is better quality control through official cooperation.
- It is based on European standards.
- The data is generated by trustworthy institutions providing unbiased products.

ESDIN: a European Spatial Data Infrastructure

ESDIN – a European Spatial Data Infrastructure with a best practice network – is dedicated to maximizing the use of data from our National Mapping and Cadastral Agencies (NMCAs). Funded by the eContentplus programme of the EU, the project promises to provide the guidance needed to meet the demands of the INSPIRE legislation.

The project emerged during an exciting period in the history of digital content. ESDIN embraces the trends we have seen developing in the use of GI and in parallel markets as the

world wakes up to the power of geographic information and insists on simpler ways to access it.

As Geographical Names is one of the themes ESDIN covers, the project offered the chance to demonstrate geographical names services as the cornerstone of a future European Spatial Data Infrastructure. As a result EuroGeoNames is referenced for the proposed extended INSPIRE specifications that ESDIN offers, and implementations within the ESDIN project include EuroGeoNames.

Exploiting expertise from the national mapping, commercial development and academic worlds, the twenty partners offer best practice when maximizing the use of public sector mapping data.

Making EGN sustainable – a European Location Framework

EuroGeographics' projects support key European initiatives designed to reduce the barriers to use and re-use of crucial public sector information. Our projects and products show that the INSPIRE dream is real and will be made more effective by applying common practical approaches to harmonization in data and services. All our work helps create a consistent and trusted European Location Framework (ELF).

A service based on official geographical names (including their exonymns) is a crucial first step toward the ELF.

EuroGeoNames continues to be supported and extended by EuroGeographics, together with the German Federal Agency for Cartography and Geodesy (BKG) and the Finnish Geodetic Institute (FGI).

As a first stage, in spring 2009, the EGN project coordination was transferred to EuroGeographics. BKG continued to function as the 'service centre' for hosting the EGN Central Service, the Reference Application and the exonyms and other variant names database, as well as for providing technical support to NMCAs and pilot customers. From mid 2012 this function, as well as the technical support activity, will be transferred to FGI.

We have been successful in growing the connections of national databases to the EGN service to 15 by the beginning of 2012 and the initiatives to grow this further include awareness campaigns and activities with key users, suppliers and influential groups. These activities include:

- Involvement with the United Nations group of experts on geographic names;
- Providing necessary advice, software and online tools to ease the task of mapping to the EGN specification and the creation of local web feature services;
- Workshops and questionnaire with national mapping and cadastral agencies on proposed pricing and licensing, access and supply preferences;
- Increasing basic coverage for the whole of Europe by including names data from alternative sources;
- A test programme, interviews, questionnaires and meetings with all key customer groups;
- Boosting the number of connected countries to 20 with an active programme aimed at connecting 27 nations;
- Incorporating into the EuroGeographics demonstration European Location Framework (ELF)⁷ to allow search and download of geographical names.

The creative problem-solving approach by the EGN Coordination Committee sets out a plan to improve accessibility and usability of the offering and now shapes the proposition for a sustainable EuroGeoNames. EGN may become the first INSPIRE-compliant service in Europe providing excellent opportunities for EuroGeographics and its members to meet future requirements in their role as reference information providers for national and European spatial data infrastructures. After 2012 the ambition is for EGN to become a component in EuroGeographics' services infrastructure.

Future plans

Events like the one-day workshop on EGN at the UNGEGN regional conference in Zagreb (11 February 2011) have helped us to engage with suppliers and users at the same time. Such

⁷ EuroGeographics' activities focus on underpinning the European Spatial Data Infrastructure with the definitive reference data collected, maintained and provided by our members. This infrastructure will provide the location framework for Europe as comprehensive resource for government, businesses and citizens. Find more about the European Location Framework (E.L.F) from the white paper:
<http://www.eurogeographics.org/sites/default/files/E.L.F%20white%20paper%20v.1.0.pdf>

activity will boost the number of connections as we gain a unique insight on core issues and take the opportunity to address them.

Our future plans are backed up by a budgeted commitment from EuroGeographics specifically to:

- Improve user interfaces to meet user requirements;
- Implement a new and more flexible and performant cloud-based architecture to accommodate different feeds and cache NMCA data;
- Launch with an offline gazetteer product;
- Continue with a phased approach to other offerings;
- Increase the exonym and variant names database content;
- Continue to improve the quality of content by providing free analysis of supplier data;
- Transfer the hosting and technical support responsibilities from BKG, Germany to FGI in Finland;
- Implement a simplified data model that is standards-compliant, removes some complexity for the suppliers and yet retains the ability to provide multilinguality and the association of exonyms to endonyms.

EGN is EuroGeographics' first example of a European Location Framework service; a framework of key geographical references from authoritative sources that help assure the reliability of services using and re-using Public Sector Information. This common reference will help services interoperate and make cross border and pan-European services far more consistent and reliable.

References

- EuroGeoNames (EGN) (2008a) *EGN Deliverable D3.3, EGN Metadata profile*. Available online at [http://www.eurogeographics.org/sites/default/files/070920_D3.3_Metadata%20profile.pdf] (accessed 30 March 2012).
- EuroGeoNames (EGN) (2008b) *EGN Deliverable D4.2e, Data model conceptual schema & documentation*. Available online at [http://www.eurogeographics.org/sites/default/files/D4_2e_EGN_Conceptual_schema_documentation_0.pdf] (accessed 30 March 2012).
- EuroGeoNames (EGN) (2008c) *EGN Deliverable D6.5, EGN Web Services profile – specifications for impementation*. Available online at [<http://www.eurogeographics.org/>]

sites/default/files/090205_D6.5_EGN%20Web%20Services%20profile%20%20specifications%20for%20implementation.pdf] (accessed 30 March 2012).

EuroGeoNames (EGN) (2009a) *EGN Deliverable D1.7, Final Report D1.7* Available online at [http://www.eurogeographics.org/sites/default/files/D1.7_EGN_Final_report.pdf] (accessed 30 March 2012).

EuroGeoNames (EGN) (2009b) *EGN Deliverable D7.4, Documentation EGN Web GIS Reference Application*. Available online at [http://www.eurogeographics.org/sites/default/files/090205_D7.4_EGN_RefApp_Documentation_1.0.pdf] (accessed 30 March 2012).

EuroGeoNames (EGN) (2009c) *EGN Deliverable D8.2, EGN ArcGIS Extension Documentation*. Available online at [http://www.eurogeographics.org/sites/default/files/090205_D8.2_EGN_ArcGIS_Extension_Documentation.pdf] (accessed 30 March 2012).

EuroGeoNames (EGN) (2009d) *EGN Deliverable D11.2, Implementation Plan*. Available online at [http://www.eurogeographics.org/sites/default/files/D11.2_EGN_Implementation_Plan.pdf] (accessed 30 March 2012).