

ECP 2005 GEO 038026 EGN

EGN

Annual Report

1st September 2006 – 15th August 2007

www.eurogeonames.com

Deliverable number/name	<i>D1.5 1st Annual Report</i>
Dissemination level	<i>Public</i>
Delivery date	<i>31 August 2007</i>
Status	<i>Final</i>
Author(s)	<i>Jörn Sievers, Pier-Giorgio Zaccheddu</i>



eContentplus

This project is funded under the *eContentplus* programme¹,
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

¹ OJ L 79, 24.3.2005, p. 1.

1 Table of content

1 TABLE OF CONTENT 2

2 PROJECT OBJECTIVES 3

3 CONSORTIUM 5

4 PROJECT RESULTS/ACHIEVEMENTS..... 6

5 TARGET USERS & THEIR NEEDS 8

6 UNDERLYING CONTENT [D2.1], [D3.1], [D3.2]..... 11

7 SUMMARY OF ACTIVITIES..... 13

8 IMPACT & SUSTAINABILITY 15

9 FURTHER INFORMATION 16

10 REFERENCES / RELATED DOCUMENTS 17

2 Project Objectives

Short description of the problem addressed by the project and its objectives.

Summarize briefly how the project contributes to the programme objectives.

The EuroGeoNames (EGN) project started successfully on 1st September 2006 and will last for 30 months until 28th February 2009. The total project budget amounts to 1,8M €. It is co-financed by the European Commission (contributing 50% of the total costs) within the eContentplus programme. EGN is coordinated by the Federal Agency for Cartography and Geodesy (BKG).

2.1 Objectives and problem addressed by the EGN project

Access to consistent and reliable multilingual geographical names is essential for a number of uses including postal services, emergency services, navigation, tourism, property purchases, the mass media and applications such as Google Earth. In all of these areas, geographic names provide one of the most important keys for geo-referencing and accessing a variety of related information. However, considering the claims of the INSPIRE legislation there is no European standard or service for accessing geographical names, rather a patchwork of heterogeneous national services that are not suitable for the European market.

The EuroGeoNames (EGN) project will address this problem by establishing a European infrastructure and services for the exploitation and re-use of existing public sector geographical names data sources held heterogeneously and distributedly across Europe. One important deliverable will be interoperable internet services that will link and provide access to the official, multilingual geographical names data, updated in a consistent way and maintained at the source level by the responsible organisations.

Within the EU-funded project duration the “critical mass of content aggregated”, as requested by the eContentplus programme, will be fulfilled by aggregating data for between 5 and 10 European countries – comprising most probably also Candidate as well as EFTA countries – by connecting their national databases in the EGN infrastructure. Currently, national mapping and cadastral agencies (NMCAs) belonging to 15 countries (Austria, Cyprus, Czech Republic, Germany, Finland, France, Hungary, Lithuania, Latvia, The Netherlands, Norway, Slovakia, Slovenia, Spain, Turkey) presently constitute the ‘EGN Reference Group’ and are the potential data suppliers having declared themselves prepared to make their national geographic names data stocks available for the requirements of the EGN project.

EGN will be targeted primarily at value added resellers (VARs) and service providers to develop specific applications for their customers and deploy value-added GIS products by using the EGN Web Service. The end user will have access to this information at least through a Web GIS reference application which will enable searching in all official European languages, including officially recognized minority languages.

Amongst other deliverables, EGN will provide a sound and robust pan-European data model for geographical names data. This will provide the basis by which to both harmonise current heterogeneity and also stimulate NMCAs towards better integration of geographical names data within their own national spatial data infrastructures (SDIs). Additionally, EGN will provide links to find and relate variant names, in particular exonyms², to the respective

endonym² and indicate how it is used officially in the country where the geographical feature is located.

Furthermore, EGN will provide a solid business model and marketing concept for the re-use and value-adding of European geographical names data as the rationale for the time beyond the EU-funded project duration.

Last but not least, EGN will also capitalise on the established network and knowledge of the community of European geographical names experts collaborating intensively in the UN Group of Experts on Geographical Names (UNGEGN), with respect to using authoritative geographical names sensitively and appropriately. Sometimes wider political issues are related to them and almost always, linguistic issues have to be taken into account, e.g. regarding questions to the adequate use of main and/or minority languages, exonyms, pronunciation, etc.

To sum up, EGN will set up a full-blown distributed multilingual geographical names data infrastructure and services for Europe, which will help to promote cultural diversity and multilingualism.

2.2 Contribution of the EGN project to the eContentplus programme

The overall aim of the eContentplus programme is to, “*facilitate access to and use of European digital content*”. The EGN project addresses a subset of Europe’s digital content – geographical names – within the targeted area of geographic information and specifically the objective of, “*improving the enabling infrastructure*” to arrive at “*cross border, interoperable and/or seamless geographic information*”. The project will result in quality content (geographical names), aggregated from national, multilingual sources, that can be accessed easily (using ‘state of the art’ technology) by service/product providers and users for a variety of applications.

Apart from the fulfilment of overall aim of the eContentplus programme, the EGN project addresses especially multilingual aspects. EGN will facilitate access to the European geographical names data (maintained at the source level by the responsible organizations). The names data will be provided equally for all languages officially spoken in the European countries, including names data maintained in the officially recognized minority languages. Thus, EGN will also promote the cultural diversity and multilingualism.

EGN will provide interoperable data and populate the European Spatial Data Infrastructure (ESDI) [Inspire Directive 2007] through its Web services. EGN will demonstrate how the public sector information [PSI Directive 2003] from European countries’ existing geographical names data can be made available through interoperable services, to be exploited at the European level by public sector authorities and private sector players.

Where applicable, EGN will deploy open standards in services and applications. Hereby, EGN will create a prerequisite for and demonstrate how European geographical names information could make a corner-stone for the production of both commercial and public value-added services (such as topography, traffic and tourism information, trans-national emergency services, location-based services at a pan-European level, environmental analysis).

The EGN project has a concrete and hands-on approach and will open up information that already exists in digital form. The problems faced today are firstly the use of different data models, data exchange formats, etc. in different countries, secondly the lack of visibility of

² Definitions: **Endonym**: Name of a geographical feature in an official or well-established language occurring in that area where the feature is located. **Exonym**: Name 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.

and accessibility to existing data resources, and lastly the lack of transparency of pricing policies and other conditions to obtain the data. The project will promptly address all problems of the first and the second level. For the last level, at least, some of them will be touched on by providing concepts for an improvement.

The EGN project will unlock and promote the effective pan-European re-use of public sector information via a state-of-the-art infrastructure and services. It will demonstrate how the important public sector names information can be accessed and aggregated on a pan-European level.

Hence, the EGN project is fully compliant with the objectives and the general approach of the eContentplus programme.

3 Consortium

Provide a brief description of the consortium members and their roles in the project.

The EGN project is carried out by an international consortium comprising nine partners from the economy, science and public administration sectors and coming from five countries (Austria, Germany, The Netherlands, Slovenia, Great Britain) as well as EuroGeographics, the association of the national mapping and cadastral agencies (NMCAs).

Additionally, the consortium partners represent a wide range of organizations combining the benefits of regional distribution with the benefits of public/private partnership:

- The Consortium partners from organisations/institutions of Germany, the Netherlands and Austria provide long-term experience with regard to the standardisation of geographical names. They are intensively collaborating within the Dutch- and German-speaking Division (DGSD) of the United Nations Group of Experts on the Standardization of Geographical Names (UNGEGN). The original concept for the EGN project was first discussed and developed in the DGSD.

The Federal Agency for Cartography and Geodesy (BKG) is the project coordinator and mainly responsible for the overall co-ordination and management of the EGN project. It acts as primary contact person to the EC, arranges contacts with other organisations and executes the technical and financial management.

The University of Utrecht (UU) mainly provides a detailed description of user and business requirements for the EGN infrastructure. This description forms the basis for the development of the EGN data model and provides also use cases for the web services' development, the applications and the business model.

The Bundesamt für Eich- und Vermessungswesen (BEV) addresses, on the one hand the iterative testing and refinement of the EGN infrastructure and services against the user/business requirements. On the other hand, it evaluates and assesses Component Interface Tests (CIT) which will exercise the server and/or client component software's ability, document and test the EGN web services in different IT environments, in order to properly implement a robust and scaleable infrastructure.

- Slovenia, through its mapping agency (SMA), is in this context representing the Slavic-speaking area and thus also serves as a bridge between the orthographic diversity of Europe. Within the EGN project SMA is mainly responsible for the evaluation and assessment of all aspects of the national repositories of geographical names data that may have an impact on the design of the EGN infrastructure.
- EuroGeographics – the Association of the European national mapping agencies, with Head Office in Paris (France) and 47 members from 41 European countries – is actively engaged

in defining the structure(s) of the reference data for a European Spatial Data Infrastructure (ESDI) and the provision of reference data specifications and guidelines for the ESDI set-up. Within the EGN project EuroGeographics ensures continuous dissemination of project progress and serves to raise awareness of the project in order to promote the EGN project amongst other potential users and to maximise usage at the earliest practicable juncture.

- The companies Geodan Software Development & Technologies B.V (The Netherlands), GeoTask GmbH (Germany) and ESRI Geoinformatik (Germany) are involved in building national and European geo- and metadata portals, GI software and system technology for both, the private and public sectors. Drawing upon experiences elsewhere, these companies will be able to expand the market potential for the (multilingual) EGN infrastructure and will assist in the evolution and expansion of a core SDI component for Europe. Geodan is mainly responsible for the technical setup of the EGN infrastructure and services. It includes the elaboration of the technical service requirements, software selection and software engineering for the implementation of the EGN central service and localised components. Geodan takes over the technical support for the web services implementation in the European countries. GeoTask provides the graphical user interface for names searches and for the visualization of the search results, including the selection, programming and documentation of the software components and interfaces. Additionally, GeoTask estimates the market potential for the EGN infrastructure by conducting a cost/benefit analysis, which provides the information to develop a draft business model and a marketing strategy for EGN. ESRI is mainly responsible for the design of an application in a commercial context. As an example this application will integrate the EGN central service into a specific client application.
- EDINA, National Data centre based in the University of Edinburgh (Scotland), provides not-for-profit services to the UK higher and further education academic teaching and research communities. It is actively engaged in establishing an SDI for the UK academic sector and provides on-line services delivering a wide range of geospatial data, support and advice, including provision of academic access to the National Mapping Agency's (Ordnance Survey Great Britain) data holdings. Within the EGN project EDINA evaluates and assesses existing European reference data models. The findings here are used as the starting point for the development and documentation of the conceptual and application schemas for EGN, ensuring synergy with other reference data that will underpin the ESDI.

4 Project Results/Achievements

In case of an **annual report**, describe major achievements (e.g. completion of market and user requirements survey, completion of demonstrator design or implementation, initial reactions and feedback from users, first results).

The establishment of the EuroGeoNames (EGN) infrastructure and associated services consists of four periods, the initiation phase (month 1 to 8), the development phase (month 9 to 17), the implementation phase (month 18 to 26) and the completion phase (month 27 to 30). Accordingly, this annual report comprises the *initiation phase* as well as the the major part of the *development phase*.

The following figure 1 shows the spiral engineering process necessary to set-up EGN. The tasks and actions which have already been completed were ticked. The tasks and actions which have been started and are currently ongoing are indicated with an exclamation mark:

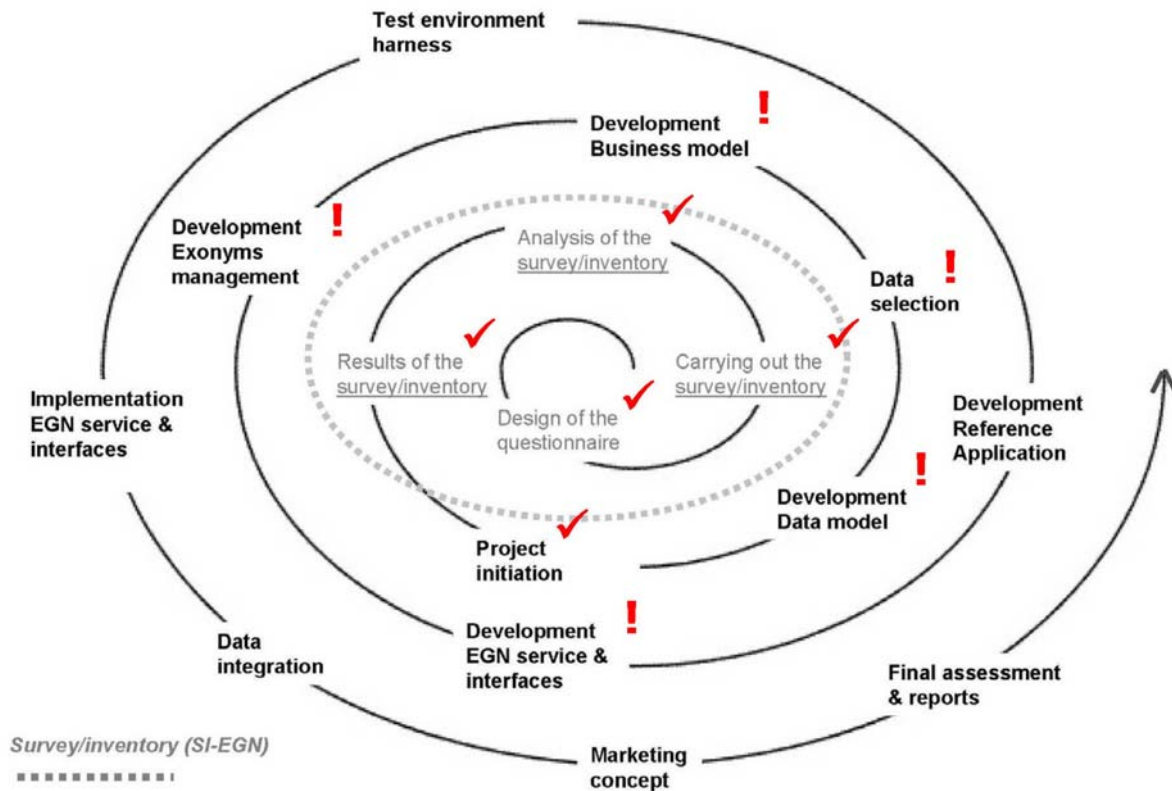


Figure 1: Spiral engineering process necessary to set-up EGN; ✓ = Tasks accomplished; ! = Tasks in progress; SI-EGN tasks (rendered in grey tone) are not part of EGN, but are a necessary prerequisite.

The first period, the *initiation phase* (month 1 to 8), was mainly characterized by the **communication** between the EGN Consortium and the EGN Reference Group. Amongst others, the aims, expectations and deliverables of the EGN project were clarified and the geographical names data to be provided by the EGN Reference Group was discussed and preliminarily evaluated and assessed. Furthermore, **key technical issues** to be **addressed** by the EGN project were identified and a general consensus were reached on the content for a common EGN conceptual schema for geographical names data in Europe.

The market potential for the EGN infrastructure in terms of potential applications and related business actors and stakeholders from both public and private sectors was estimated by conducting a **cost/benefit analysis**. This market analysis was the rationale for providing a detailed description of **user and business requirements** (see chapter 5).

The outcomes as well as all other deliverables of status ‘public’ are available through the EGN website: www.eurogeonames.com

Within the second period, the first part of the *development phase* (month 9 to 12), two major efforts the EGN Consortium tackled can be identified. On the one hand the discussion and development of the **EGN conceptual schema**, against which the national data models have to be converted to. On the other hand, the endeavour of developing a **concept for a EGN web services architecture** was supported by the EGN Consortium (incl. data access security issues as well as the documentation of web services requirements).

All dissemination activities led to a **dissemination plan** covering the project lifetime. It will be regularly expanded and maintained as part of the dissemination activity. The already existing EGN section in the EuroGeographics website was extended. To raise awareness on the EGN project a **flyer/brochure** was created, describing the project for general communication (more simple and for general purpose). It is publicly available through the EGN website (see above).

The interdependencies between EuroGeographics, the association of the European National mapping and Cadastral agencies, the INSPIRE initiative of the European Union and the UNGEGN were safeguarded through the involvement of these groups in document review processes or the invitation of members to specific **meetings and workshops**, five of them having been organized since the project start: in Frankfurt am Main in October 2006, in Utrecht in January 2007, in Madrid in March 2007, in Frankfurt am Main in May 2007, in Praha (Prague) in May 2007 and in Dortmund in June 2007.

The objectives and outcomes of the meetings and workshops are available through the EGN website: www.eurogeonames.com

As one of the major achievements it should be stated here, that EGN brings together complementary expertise existing in different organisations through the EGN Consortium, the ‘EGN Reference Group’ – as being the data providers – and the ‘EGN Group of Interest’. The latter group is comprised of stakeholders embracing the full ‘value chain’ from data providers, technology partners, and value-added service applications. The group can join, comment and shape the project activities from the beginning and provide ‘**customer feedback**’ by participating in workshops and the **online user forum**. The online user forum was developed and was launched in March 2007. It is accessible through the EuroGeoNames website for the purposes of information dissemination and the publication of project progress and results to all stakeholders interested in the EGN project. In addition, all members of EuroGeographics receive information about the project and have the opportunity to contribute to its results through the existing network of Expert Groups, workshops and other communication channels. By that, EGN demonstrates tight co-operation between both public and private organisations in Europe, which increases the basis for the commercial services for a multilingual and multicultural user environment.

5 Target Users & their Needs

Provide a summary of your user groups, research on user/customer requirements and the resulting user profiles, and how the project results will respond to these requirements/profiles.

5.1 User groups, needs and requirements [D2.1]

Within the *initiation phase* of 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.

The following section summarises how names data are collected and how names files need constant updating and maintenance in the European countries. It then distinguishes between different **use functions** and different **user groups** and determines specific **user group needs**.

These are systematised in **use cases**. The ensuing needs and requirements are put in a market perspective and then followed up by an inventory and analysis of current commercial applications of geographical names. As the main objective of EGN is to create a web-based geographical names infrastructure and services, the scope and attributes of **current free web-based names servers** were analysed, too. Finally an **information model** was developed based on the use cases mentioned above.

The following user groups and use functions have been discerned, that should be reflected in the use cases described (see table 1). The geographical names enrichment function discerned was seen as up to the value-added resellers (VAR).

functions/ use groups	normalisation = look up	translation	indexing	geocoding ~ geoparse	geo- indexing	reverse look up
finance, insurance	xxx	xx				
web sales, tourism	xxx 1	xxx 1	xxx	xx	xx 1	
marketing, health	xxx 9	x 9		xxx 9	xxx	
media	xxx 8	x 8	xx			
distribution	xx	xxx	xx	x	xx	
spatial planning	xx 7	xx		x 7	x	
map data production	xxx	xxx	xx	x	x	xx 11
emergency services, lbs	xxx	x	x 4	xxx 4	x	xxx 10
science	xxx 5	xxx 5		xxx 5	x 6	x 6
individual users	xxx 2, 3	xxx 2, 3	x	xxx		

Table 1: Functions for the use groups discerned, on the basis of needs (xxx), wants (xx), nice to have (x). Numbers refer to the use cases (highlighted). The following use cases were described: 1) hotel booking use case, 2) EGN name server use case, 3) geoportal map application use case, 4) emergency map use case, 5) geoparsing use case, 6) historical research use case, 7) metadata search use case, 8) geonames checker use case, 9) real estate use case, 10) coordinate emergency use case, 11) query names in bounding box. The last three rows have been added and are not based on the original questionnaire used by Geodan.

From the use cases, the following requirements for the EGN infrastructure and services emerged:

Required in terms of

content: all place names and their variant names in all European languages, their gender, number and their coordinates.
-All major lakes and rivers, mountains, physical and administrative areas, all named landscape features. Named areas should have polygons. Previous, replaced names should still be in the database.
-Finance and Insurance and Use case 4) and 9) also require street addresses and major points of interest.
-Use case 8) also requires pronunciation information.
functionality: at place name input show coordinates or bounding boxes (and show adjacent names), show feature types, show variant names, fuzzy name search (also: sounds as:); show name on a map (use case 2), combined query (feature and name or feature and area or feature and coordinates).
Use case 8 requires a pronunciation facility. Proximity analysis for use cases 1) and 6).
user interface (UI): ease of use is crucial. Purpose of UI elements must be immediately clear. Multilanguage is required.
integration: available 24/7, protocol agnostic but ADL or OGC preferred. Ability to perform multiple iterative searches.
data quality: names data should be annually updated. Spatial accuracy from 10 (use case 4) up to 100m (use case 3).

Box 1: EGN infrastructure and services requirements

A study of the geospatial information marketing possibilities identified unexpected market players such as insurance companies and health research. Furthermore, commercial users wanted to use wide-spread software technology, they wanted straightforward incorporation of the gazetteer service in their existing infrastructure and they wanted an easy exchange of information, both through import/export and through open service capabilities.

In the inventory, it was registered whether operators were from the public or private sector, their business model and platform was indicated, service providers and search categories used: place names (9), administrative units (7), addresses (4), postal codes (3), motorways (2), other areas (2).

Most common search categories in current commercial applications were complete address information and administrative area searches. The accuracy required was point coordinates for address information and bounding boxes for all other name categories. Annual updating of information seemed to be the norm.

5.2 Information model [D2.1]

One objective of the *initiation phase* of EGN was to elaborate an information model consisting of geographical feature categories that should be included and the attribute information deemed necessary for each category provided by the EGN infrastructure and services. The information model was the rationale for developing the EGN conceptual schema for geographical names data.

Based on the requirements in box 1 (taken from the use cases) 5 main name categories are discerned (one of which, the man-made feature category, is further subdivided), and pronunciation is added as a name attribute.

Based on the user interface requirements in box 1, language should be an attribute, as well as the date a name was last changed.

Street addresses and postal codes came up as a requirement in two use cases as well as in the market analysis and in the commercial applications. In view of the fact that they were beyond the scope of EGN, and also because all of the competing geographical names services did not provide street addresses and postal codes either (see chapter 6.2), it was decided not to require street names/addresses for the time being. It should be discussed whether to leave room for extending the EGN functionality with street address matching capabilities in future, so that EGN would be able to compete with this address matching possibility in route planners.

In view of the fact that after street address matching the second most frequently used search category was administrative area search, it was decided to include hierarchical features in the information model.

The most extensive competing geographical names services all had identifiers (IDs), either for the topographical features named or for the geographical names. They also had information on the status of the names and the dates this status changed. As the official character of EGN is stressed, information on this aspect of the names should be made available.

As it was decided not to go into data enrichment and leave that to value-added resellers, the present attributes are rather few. The metadata for the topographical features would refer to data quality aspects such as up-to-dateness, completeness and geometrical accuracy.

EGN Information model

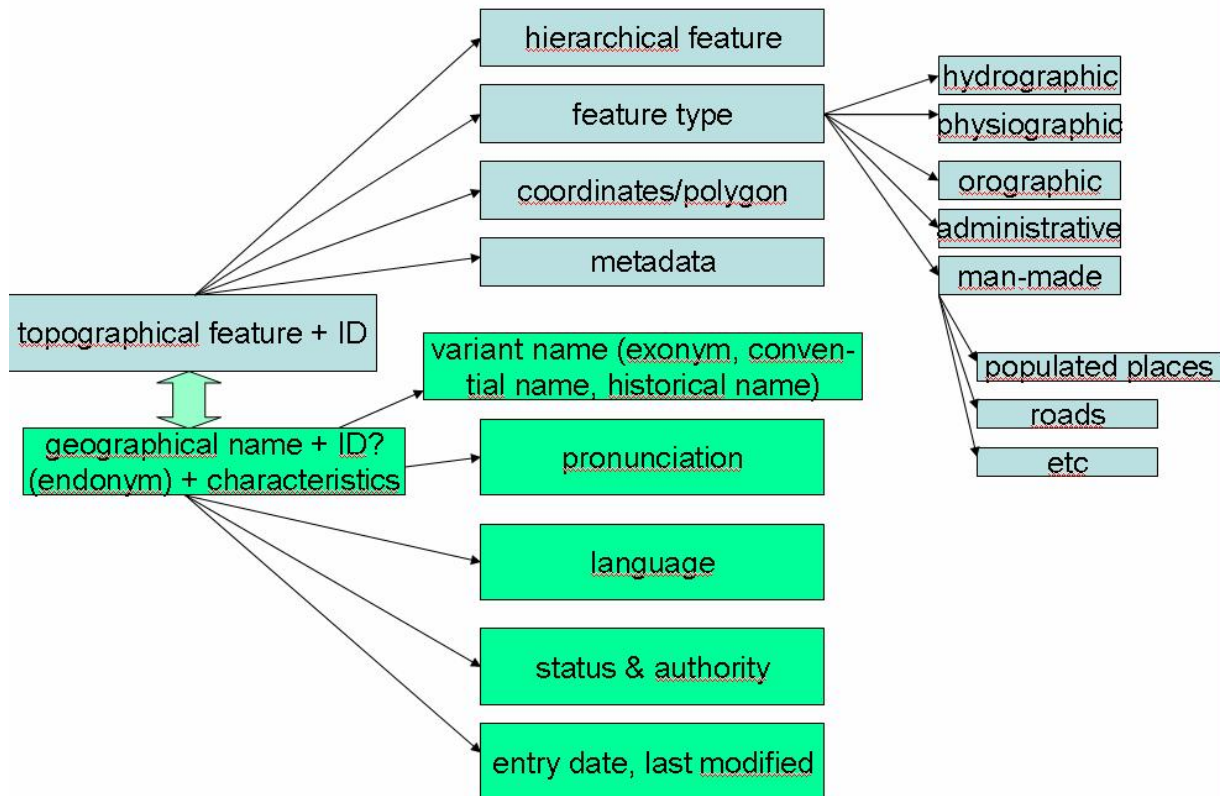


Figure 2: Proposed EGN Information model, based on use cases and market analysis

6 Underlying Content [D2.1], [D3.1], [D3.2]

The quality and quantity of the digital content (and related metadata) contributed to the project, as well as the criteria for its selection, must be clearly identified.

6.1 Underlying content in terms of quantity

The quantity of the geographical names data in the respective countries is quite different 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. Within the *initiation phase* of EGN remarkable time was required for analyzing the data to be aggregated during the project.

Information on available data had already been gathered during the SI-EGN phase [SI-EGN 2005]. Since then contact to the EGN Reference Group has been maintained and consolidated. As an outcome further details of the data were compiled and documented as an important source of information, which had relevance to large parts of the project, e.g. development of the EGN conceptual schema, the EGN services architecture and interfaces as well as for determining priorities for the web services implementations. The documentation contains information as provided by the Reference Group members on:

Feature types

Main feature categories are displayed for which geographical names data are available: administrative units, hydrographical features, transport features, populated places, orographic features and relief forms, other features.

Geographical names attributes

Indicates what kind of attribute information are linked to the geographical names (the name might be linked to the attribute information, or in topographic database would be part of the attribute information).

Map scale indicators

Shows the scale of the source document, or the resolution of the database of features to which the geographical names are linked.

Metadata

Information about availability of metadata for geographical names and compliance to metadata standards.

Codepage

Information about alphabetic character sets (codepages) used by each EGN Reference Group member with regard to the languages supported by the operating systems.

Exonyms

Information about the existence of a list of exonyms maintained by each EGN Reference Group member.

In summary, at this stage only a rough estimation can be given with regard to the **quantity** of the data that is going to be aggregated as provided by each country. Considering that in the primary phase of EGN data selection will be focused to small/medium scales (1:1 million, 1:250,000, 1:100,000), the quantity of geographical names stored and maintained in the countries vary from 50,000 to millions of entries.

From the technical point of view, the names data are either stored in vector databases as attributes to the topographic features (approximately 50 Gigabyte per country) or in single stand-alone databases (approximately 0,5 Gigabyte per country).

Thus, assuming that all fifteen countries mentioned above will be connected with EGN, the volume of the data that will be brought into the project is, in total, about 550 Gigabyte and the quantity of geographical names lie in between 2,8 million and 3,5 million entries³.

Within the *development phase* of EGN a metadata profile has been proposed. This metadata profile will be referred to the “dataset” as a whole, i.e. in Germany to the German geographical names dataset GN-DE. This metadata profile will be a list of elements/attributes to be filled out by each EGN Reference group member (NMCA) according to ISO19115 and the requirements stated by the Inspire Drafting Teams. Thus, within the EU-funded project duration of 30 months 15 NMCAs will fill out this profile, which will then be stored and maintained in the metadata system provided by EuroGeographics. Each profile will contain information about the “dataset character set used”, “dataset responsible party”, “keywords”, “dataset language(s)” an “abstract” etc.

6.2 Underlying content in terms of quality

A survey of current free commercial geographical names services on the web shows seven providers (Geonames, GEOnet Names Server (GNS), Getty thesaurus of geographical names, the Fuzzy Gazetteer, the ADL Gazetteer service, Earth Search and World gazetteer).

Apart from Fuzzy Gazetteer all of them use the names taken from the GNS, maintained by the US National Geospatial Intelligence Agency (NGA) and the US Board on Geographical Names. The Fuzzy Gazetteer was maintained by the European Commission and the European Joint Research Centre.

³ The quantity of geographical names is based on the information provided by the countries themselves within the survey on European geographical names data (questionnaire replies). The final report of the survey (SI-EGN final report) was attached in the EGN proposal submitted in November 2005 and is also available through the website: www.eurogeonames.com

Attribute data all these services had in common are names, coordinates, feature codes, country codes. The functionality they all shared was the location of the named objects on the map. Half of them worked with unique name or object/feature identifiers (IDs).

The **unique selling points** (usp) of the EGN infrastructure and services for Europe and its **quality** are that,

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

The current geographical names web services reviewed score as follows in comparison with EGN regarding these usp's (see table 2):

Unique selling points	EGN	current geoname services	current route planners
primary data	yes	no	?
official data	yes	?	partly
high quality data**	yes	?	yes
up-to-date data	yes	no	yes
complete coverage	not yet*	no	depends on category
according to European standards	yes	?	yes
including street level data	no	no	yes
coverage	Europe	whole world	Europe +

Table 2: Comparison between EGN infrastructure and services, current name servers and route planners. *) complete coverage is envisaged. **) comparable to 1:50 000, coordinates to 10m

7 Summary of Activities

Draft a short section for each substantial area of work completed/started, tailored to reader needs rather than revolving around work packages.

The **annual** report should also describe where the project is 'positioned' for the next year, describing which are the activities that will be completed and the results that are expected.

7.1 Status-quo of the EGN project

EGN user and business requirements [D2.1]:

A comprehensive study on user and business requirements for geographical names data was conducted and a report was produced. The report makes available sections on the maintenance of geographical names files, on the needs of geographical names user groups, on data sourcing for geographical names, on market analysis for geo-information and geographical names, on current commercial applications of geographical names as well as on the characteristics of geographical names services available now. It ends with the definition of an information model which was the basis for several tasks to be conducted within the *initiation and development phase* of the EGN infrastructure and services.

EGN Data model conceptual schema [D4.2]:

This document presents the EGN Conceptual schema (a map of concepts and their relationships). It represents the result of an iterative process of development under the auspices of the EGN Project. It has been produced by the EGN Consortia in league with its wider EGN Reference Group (data providers) and within the context of INSPIRE best practice guidance. This deliverable will form the basis for the development of an Application profile/schema. The Conceptual schema should be considered informative rather than definitive at time of writing as subsequent developments within INSPIRE and during implementation may necessitate model revision. Given the broad audience this document is aimed at, a tabular version of the schema as well as simplified and extended UML variants were produced for the purposes of reaching as broad an audience as possible.

EGN infrastructure and services requirements [D6.1]:

A set of functions that the EGN infrastructure and services should support was elaborated and defined. These functions cover the technical system requirements based upon the defined user and business requirements. It is proposed that the EGN web services will make use of the Web Feature Service (WFS) specification as defined by the Open Geospatial Consortium (OGC). This specification allows for all proposed EGN functions.

At each National Mapping and Cadastral Agency (NMCA) a WFS will be deployed. Those local services will be accessible through a central service (also WFS). Two scenarios are proposed for deployment and maintenance of local services: In the first scenario, NMCAs are fully responsible for building and maintaining the service, according to EGN specifications. In the second scenario, the EGN consortium will implement and deploy the necessary software. The purpose of the EGN infrastructure is to provide information on geographical features in response to queries that use a toponym (an endonym, variant or exonym). Based on wishes and demands from the stakeholders in the EGN project, a set of functions that the service should provide is defined. This functional design is presented as well as ideas on how to implement such a service.

Together, the functional design and services architecture can be seen as a fulfilment of the system requirements. The aim of the EGN project is to try to fulfil those requirements as much as possible.

The functional design and services architecture outlined will be used as a basis for developing a prototype of the EGN web services that can be used for testing as well as the development of clients making use of the EGN infrastructure.

EGN business and pricing model as well as marketing strategy [D10.1]:

Within the *initiation phase* of EGN a cost/benefit analysis was carried out based on the project members' expertise in the field of market analysis for web Services and geographic information. It defines the analysis approach, it presents a monetary assessment of costs and revenues as well as it incorporates the results of the user requirements and business model studies. The analysis was based on a survey which was carried out with potential end-users (companies and governmental institutions) and value-added resellers (data- and service providers as well as system integrators and solution providers).

Within the *development phase* of EGN a draft for a potential business model was prepared based on the cost/benefit analysis. A pricing scheme as proposed in the cost/benefit analysis still has to be discussed and agreed upon with the EGN Reference Group members (NMCAs). It is noteworthy that the NMCAs – and the EGN Consortium– are comprehended as trustworthy institutions providing unbiased products.

The main points to note as an outcome of this substantial area so far is, that it is important to motivate more NMCAs to become EGN Reference Group members and to give them an incentive through the envisioned revenue share.

7.2 Next steps, activities and results expected for Year 2

The next year of the EGN project will comprise the last part of the *development phase* (month 13 to 17) and the first part of the *implementation phase* (18 to 24).

It will be focussed mainly on the **development and testing of the software components**, the web services prototypes and interfaces and the application prototypes.

A major effort will rely on the **implementation plan** for subsequently phase in the national databases of the 15 EGN Reference Group members across the project duration as well as on facilitating easy and rapid **linkage of inquired variant names (exonyms)** with their corresponding endonyms and vice versa.

Last but not least, especially the **agreement on a solid business model and marketing concept** for the re-use and value-adding of European geographical names data for the time beyond the EU-funded project duration will be tackled.

8 Impact & Sustainability

This section should highlight the European dimension of the project and why and how it has an impact on the target market (*e.g. due to the critical mass of content aggregated, the improved accessibility of the content*), including information on market prospects (*provide a summary of your market research, giving an overview of the market situation, future trends in your target markets.*)

8.1 Market prospects of the EGN project

The results of the market analysis indicate a high level of interest and needs, as well as a simple structure and cost model of services wanted. In short all **market prospects** agree to the communicated service-types and service levels. The proposed draft pricing model has still to be discussed and agreed upon with the NMCA's.

Although the number of use cases is high, the required functionality of the services is relatively simple and the same over all usage-types.

Common statements resulting from the market analysis are that EGN should:

- offer standardized services, at least in two types (free of charge, commercial);
- be online, especially to be integrated in prospects own applications, portals and/ or processes;
- be offline, in a format that can easily be read;
- Define and offer payment and billing mechanisms;
- Field and market the services as soon as possible.

The analysis shows that a break-even can be reached in a very short time, and that the EGN infrastructure is generating a reasonable input to expand the range of services and strengthen the implementation of a Spatial Data and Services Infrastructure in Europe.

It is important to stress that EGN refers to the overall addressable market with a market penetration of 100% - which is of course not possible. As a common measure usually 10-20% of the overall market is seen as accessible for a single player (market share). Competition has to be taken into consideration here.

8.2 Sustainability of the EGN project

The agreed understanding for the EuroGeoNames project is – comprising the hosting and operating of the web services infrastructure for EGN – to be supported and extended by EuroGeographics together with the German Federal Agency for Cartography and Geodesy (BKG) beyond the end of the project. The sustainability of the EGN project beyond the end of the Community funding will be ensured in a number of ways. These include, amongst others:

Sustainability starts with adequate and careful preparation in advance of the start of the project. The EGN project represents the conclusion of a survey [SI-EGN 2005] during which the technical and some business aspects of delivering a EGN infrastructure and services were investigated. This survey as well as all recently conducted market analyses showed that not only was it technically possible to build a geographical names infrastructure and services but also that there was clear demand for it. This already provides confidence that the results of the EGN project will be sustainable.

Central to achieving sustainability is to bring together the **key stakeholders in a partnership** of complementary skills and roles. Already, many of the key stakeholders were brought together during the survey stage of EGN and this has been carried forward as the Consortium for this project. The Consortium brings together the complete ‘value chain’ from data providers, through system developers, service delivery, value adding and distribution to end users. All partners in the Consortium, therefore, are used to working with each other, have complementary skills and a shared vision of what EGN will deliver.

The **methodology and approach** proposed in the project is also key to sustainability. The ‘spiral engineering’ approach ensures that the EGN infrastructure to be developed will be based on a clear understanding of user and business requirements and involvement of the key stakeholders.

Dissemination and awareness activities combined with an appropriate **organisational framework**, including the establishment of various ‘**user groups**’ - the ‘EGN Reference Group’, the ‘EGN Group of Interest’ and technically, the ‘Online-user forum’ - ensure appropriate involvement of the key stakeholders and the wider GI community during the project and ensure their interest in and involvement beyond the end of the EU-funded project period.

The EGN project is **accelerating standardization** of geographical names data between all participating European countries, by making a substantial contribution to the creation of standards for the storing and maintaining of homogeneous, up-to-date, and quality assured geographical names data on a pan-European level that can be used by both the public and private sector. The EGN project strictly observes and follows the claims made by the INSPIRE legislation to set up a European Spatial Data Infrastructure.

9 Further Information

If relevant, list here whatever else you deem necessary/appropriate for the understanding of the work you have done, the results you have achieved and/or the objectives you have reached.

- **Project website & online user forum:** www.eurogeonames.com

The extended project website provides project overviews and highlights; up-to-date information on intermediate and project results; project events, including e.g. meetings, conferences and workshops; contact details etc.

An online user forum was developed and launched in March 2007. It is accessible through the EuroGeoNames website for the purposes of information dissemination and the publication of project progress and results to all stakeholders interested in the EuroGeoNames project.

- **EGN flyer/brochure**

To raise further awareness on the EGN project a new EGN flyer/brochure was created, describing the project for general communication (more simple and for general purpose). It is available through the project website: www.eurogeonames.com

10 References / Related Documents

[DoW 2006] EuroGeoNames - Developing a European geographical names infrastructure and services, Description of Work, ECP 2005 GEO 038026 EGN

[D2.1] EGN Deliverable D2.1, Executive summary of final report on user/business requirements, 2007 (www.eurogeonames.eu)

[D3.1] EGN Deliverable D3.1, Tables of quantity, 2007 (www.eurogeonames.eu)

[D3.2] EGN Deliverable D3.2, Data selection evaluation and assessment report, 2007

[D4.2] EGN Deliverable D4.2, Data model conceptual schema & documentation (www.eurogeonames.eu)

[D6.1] EGN Deliverable D6.1, Web services requirements, 2007 (www.eurogeonames.eu)

[D10.1] EGN Deliverable D10.1, Cost/benefit analysis, 2007 (www.eurogeonames.eu)

[Inspire Directive 2007] 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)

[PSI Directive 2003] Directive 2003/98/EC of the European Parliament and of the Council on the re-use and commercial exploitation of public sector documents (PSI Directive)

[SI-EGN 2005] EuroGeoNames - Phase One, Survey/inventory on the state of the art of European geographical names data sources and assessment of a future European geographical names data infrastructure (SI-EGN), 2005 (www.eurogeonames.eu)

[UNGEEN Bulletin 33/2007], United Nations Group of Experts on Geographical Names, Newsletter No. 33, 2007 (<http://unstats.un.org/unsd/geoinfo>)

Project website: www.eurogeonames.eu