

ECP 2005 GEO 038026 EGN

EGN

Web services requirement documentation

Deliverable number	<i>D6.1</i>
Dissemination level	<i>Public</i>
Delivery date	<i>31 May 2007</i>
Status	<i>Final</i>
Author(s)	<i>Anne Blankert, Frans Knibbe (Geodan)</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.

Table of content:

1	Introduction	3
1.1	Purpose of this document	3
1.2	Summary	3
2	Functional design.....	4
2.1	Some definitions	4
2.2	Input and output.....	4
2.2.1	Input	4
2.2.2	Output.....	5
2.2.3	Notes.....	5
2.3	Examples	5
3	Technical architecture and data flow	7
3.1	EGN web service specifications.....	7
3.2	Implementation of a EGN Local Service	8
3.2.1	Components.....	8
3.2.2	Responsibilities	9
3.3	Compliance testing	9
3.4	Availability.....	9
3.5	User management and access control.....	9
4	References	11
	Appendix A: EGN service requirements.....	12
A.1	From the perspective of the end user	12
A.2	From the perspective of the VAR	13
A.3	From the perspective of the data provider.....	13
A.4	From the perspective of the EC.....	14

1 Introduction

1.1 Purpose of this document

In the EuroGeoNames (EGN) project, work package 6 comprises the design and development of the EGN Web Service (see the EGN Description of Work¹). This document is the first deliverable of work package 6 (D6.1). It describes the functional and technical design of the EGN Web Service.

The purpose of the EGN Web Service is to provide information on geographical features in response to queries that use a toponym (an endonym or exonym). Based on wishes and demands from the stakeholders in the EGN project (appendix A), a set of functions that the service should provide is defined. This functional design is presented in chapter 2: 'Functional design'. Ideas on how to implement such a service are presented in chapter 3: 'Technical architecture and data flow'.

Together, the functional design and service architecture can be seen as a fulfilment of the system requirements. The aim of the EGN project team is to try to fulfil those requirements as much as possible. However, it could be that some requirements can not be fulfilled. This could be the case if fulfilling a functional requirement would take up too many resources or if there a conflicts between requirements.

The designs outlined in this document will be used as a basis for further work in work package 6. Deliverable 6.2 will be a prototype of the EGN Web Service that can be used for testing and the development of clients making use of the EGN Web Service.

1.2 Summary

A set of functions that the EGN Web Service should support is defined (paragraph 2.2). These functions should cover the system requirements (appendix A). It is proposed that the EGN Web Service 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, NMCA's 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.

2 Functional design

In this chapter, the functional requirements of the EGN web service are listed. The functional requirements are based on the system requirements (appendix A).

2.1 Some definitions

To avoid any possible misinterpretations, some of the terms used in the context of describing the functions of the EGN web service are listed below:

1. A **feature** is a model of a real-world geographical object. Those objects may or may not have physical manifestations. Examples:
 - the Alps
 - India
 - the Donau
 - the Lake District
 - Paris
2. Features can be said to be of a certain **feature type**. One feature type has a fixed set of attributes and possible operations. For example, a transportation model might define the feature types 'road', 'waterway', 'railway' and 'air route'. EGN will use only one feature type: a named object. A named object is a feature with a distinguishing and generally accepted name. This means an object like 'the large tree next to my house' falls outside of the scope of EGN. Since all EGN features have the same feature type, all features have the same set of attributes and all EGN operations can be applied to all features.
3. EGN named objects can belong to one or more **feature classes**. A feature class is an element of a classification. One or more feature classifications can be used in the EGN web services. Examples of classifications are the Alexandria Digital Library (ADL) and Euro Regional Map (ERM) classifications. Examples of feature classes are:
 - lake
 - mountain
 - mountain range
 - country
 - populated place

2.2 Input and output

The service can be described by the specification of information that it needs (input) and information that is returned (output), as is done below.

2.2.1 Input

Feature selection

1. A search term (a string of Unicode characters) or a EGN feature ID.
2. Optional: A specification of how the term should be used (literal, fuzzy or 'sounds like').
3. Optional: A spatial operator (requires use of the EGN feature ID).

4. Optional: A specification of a set of feature classes.
5. Optional: A rectangle (two diagonally opposite ETRS89* co-ordinates).

Output modification

1. Optional: The desired language of the response (if not specified English will be used).
2. Optional: The set of desired feature attributes (if not specified, all available attributes will be returned).
3. Optional: A specification of geometry generalisation level (if not specified, the highest generalisation level will be used).
4. Optional: Result page number (if not specified, the first page will be returned) and maximum number of features per page (the default and maximum value is 50).

2.2.2 Output

A set of features that match the search criteria.

For each feature:

1. The EGN feature ID.
2. The values of the selected attributes.

2.2.3 Notes

1. The EGN feature ID is an identifier that uniquely identifies a feature within the EGN system (It is currently foreseen that this ID may consist of multiple parts).
2. Except for name, feature class and rectangle, none of the feature attributes can be used in a query.
3. The service will not sort output, this is left to the client application.
4. The optional page number and maximum number of features parameters enable the service to use output pagination. Pagination could be used to limit the data that can be gathered by free individual searches. Although results will not be sorted, a request with a page number specification will always yield the same results (as long as the source data have not been changed).
5. The output described above is not necessarily the information that will be presented in an end-user client application. Such an application may hide information or do some further processing.
6. Using EGN feature ID's in the query instead of a search for a name makes it possible to construct iterative searches: A first query can be build with a search term. This query returns a set of feature ID's. Those ID's can then be used for a second query.

2.3 Examples

1. A user enters the search term 'm*', specifies the feature class 'city' and requests all the names in his/her language. The EGN web service returns a list of all cities starting with the letter M.

* The European Terrestrial Reference System 1989 (ETRS89) is the standard EU coordinate system.

2. The same user uses the Feature ID to request the centroid of Madrid from the Spanish EGN service.
3. The same user uses the centroid to request all lakes within a radius of 30 kilometres of Madrid.
4. A user requests the first page of all the features in Europe by drawing a rectangle. The EGN web service returns ten features contained within the rectangle.

3 Technical architecture and data flow

3.1 EGN web service specifications

Given the functional requirements, the following choices are made for the realization of the EGN web services:

1. The EGN service will be a publicly accessible HTTP (and/or HTTPS) service.
2. All EGN web services will be non-transactional (read-only) OGC Web Feature Service (WFS), following version 1.1.0 of this specification².
3. All EGN web services will try to follow the OGC candidate implementation specification for gazetteer services³. If necessary, suggestions for change will be communicated to the OGC.
4. EGN web services will support both GET and POST methods of HTTP.

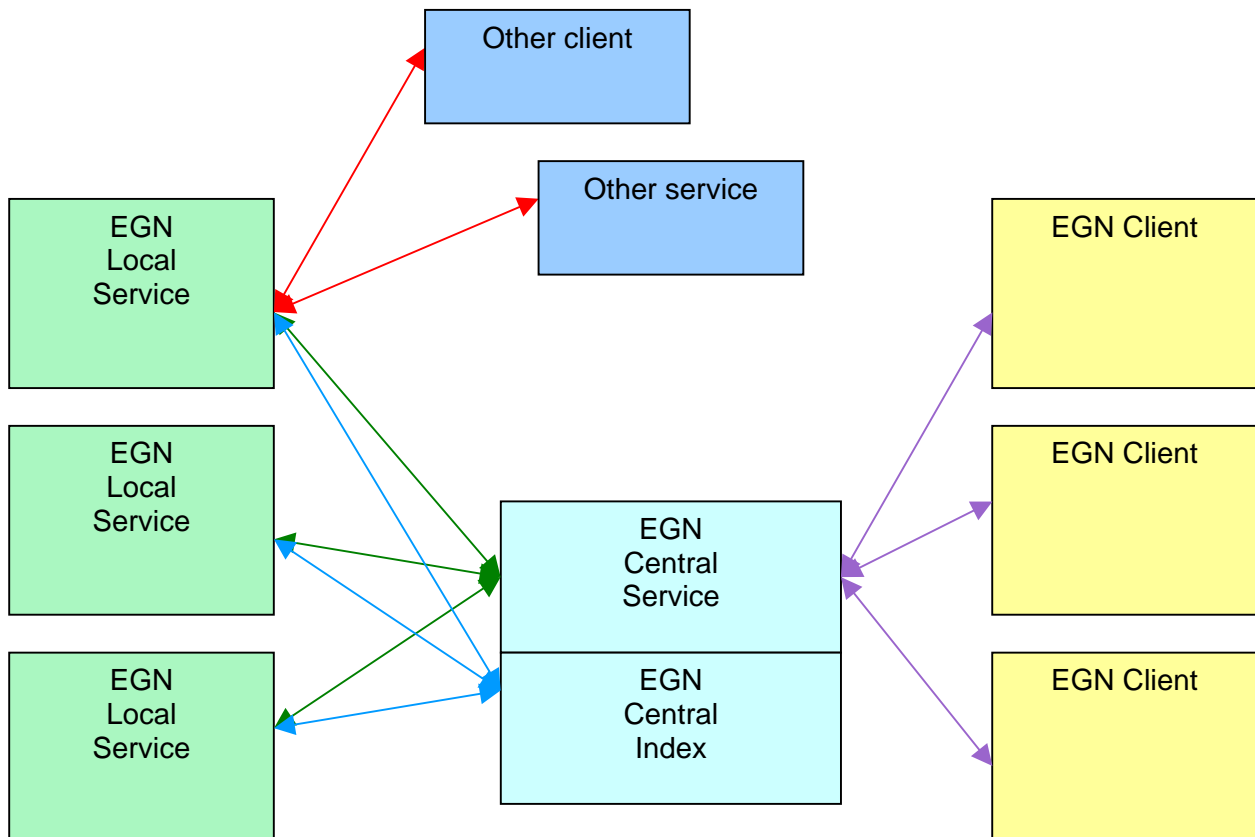


Diagram 1: All connecting arrows represent WFS connections.

Diagram 1 shows the ways in which WFS services will be used:

1. As a means of getting data (names and ID's of all exposed features) to the EGN Central Index (blue lines)
2. Clients can query the EGN Central Service with search term queries (purple lines). The EGN Central Index will be used to find the set of EGN Local Services to contact.

The EGN Central Service will query those services (green lines), process the results and send them back to the client (purple lines).

3. EGN local services can choose to provide extra functionality to users or might want to participate in non-EGN systems or SOA's*. This is indicated by the red arrows.

The EGN web service will support the following standard WFS operations:

1. **GetCapabilities:** Provides a capabilities document. The capabilities describe what a client can expect from the server. The capabilities document will identify the services as being EGN gazetteer services.
2. **DescribeFeatureType:** Provides a description of the structure of the EGN feature type. This description can be regarded as a representation of the EGN data model (EGN deliverables 4.2 and 4.3).
3. **GetFeature:** Returns a selection of feature attributes, encoded in GML, based on the specifics of the query that is received.

3.2 Implementation of a EGN Local Service

3.2.1 Components

There are two possibilities for the implementation of a EGN Local Service:

Scenario A: The NMCA will build and maintain the EGN Local Service according the EGN interface specifications.

Scenario B: The EGN partners will deliver and install the software necessary to create an EGN Local Service. In this case, the set-up will be as it is presented in the diagram below:

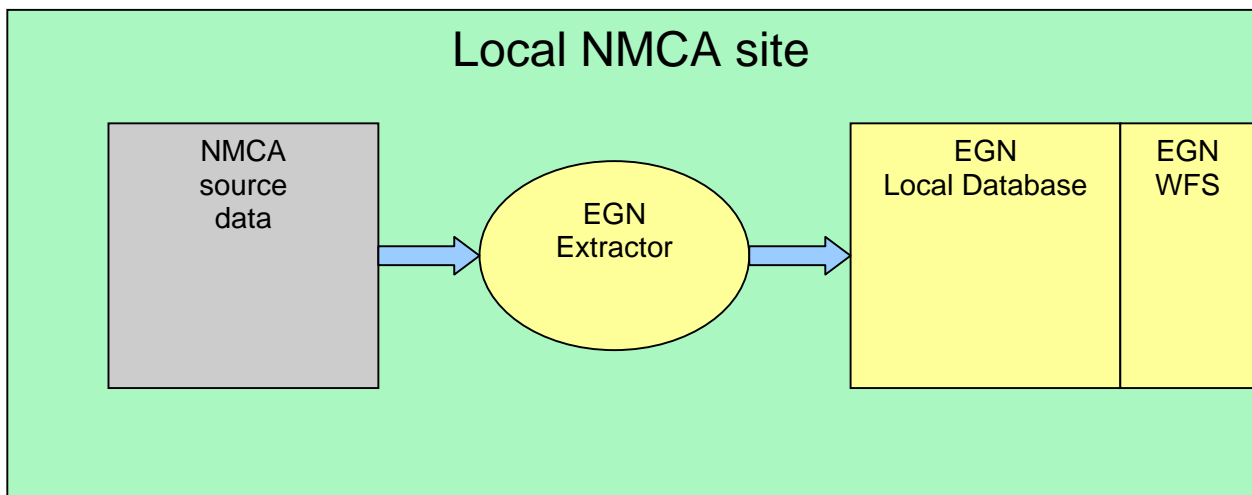


Diagram 2: Components of a EGN local service in the case the service is build by the EGN partners.

The following components can be distinguished:

* A Service Orientated Architecture (SOA) is term for loosely coupled and interoperable web services.

- **NMCA data source:** This the collection of NMCA data that is selected to be used for EGN. It can be stored in any format (RDBMS, GIS files, XML, ..).
- **EGN extractor:** A piece of software that will extract data from the data source and convert it to the EGN data model. The EGN extractor needs specifications of how to access the source data and how those data are mapped to the EGN data model. The extraction process is initiated by the NMCA, for example after each revision of the source data.
- **EGN Local Database:** An open source RDBMS (e.g. PostgreSQL, MySQL) that contains all the data necessary to allow the EGN Feature Service to respond to queries. Data are stored according to the EGN data model.
- **EGN Web Feature Service:** An open source web feature service (e.g. Deegree or GeoServer) that will be configured to be compliant to OGC WFS 1.1.0.

3.2.2 Responsibilities

In scenario A, the NMCA is fully responsible for building and maintaining the system, for keeping the system secure and for ensuring adequate responsiveness.

In scenario B, the EGN consortium will configure a dedicated EGN host computer that is installed at the NMCA site. If wanted, EGN will do the initial configuration on site. The NMCA will provide the following:

1. hardware
2. operating system (the type will be specified by EGN)
3. a connection to the internet (with a minimum bandwidth still to be specified)
4. for basic support, remote access to the system

3.3 Compliance testing

Whether or not an EGN web service instance is built and maintained by the NMCA or by the EGN partners, it is necessary to check if the service is functioning as required. For this purpose a **compliance test client** will be built. This test client can be used to query EGN WFS instances and validate the responses against the EGN interface specifications.

3.4 Availability

The EGN system will continue to work as long as the EGN central service is up. If the EGN central service can not communicate with an EGN local service or if an EGN local service takes too long to respond, results from the EGN central service can be incomplete.

Exact requirements for availability of the EGN central service and the EGN local services will be formulated in the EGN business model. Until then 'best effort' is assumed.

3.5 User management and access control

The functionality of the EGN web service will depend on the person or software making use of the service. A guest user might have permission to receive one page with results per day, while a paying organisation may have unlimited access to all information.

In the future the EGN web services will be accessible through a user management software layer that authenticates and authorizes users. An EGN transaction will be allowed or

disallowed based on the identification of the users and the content of request (guests only one page). Different usage policies may be defined for separate user roles. For registered users statistics may be kept.

The extent in which EGN can make use of standards like XACML and SAML for access control will need to be investigated. Also, the extent in which use can be made of freely available software to implement user management and access control will need to be investigated. These investigations can be performed in parallel to the development of the EGN Web Service prototype. The prototype will only support a single access policy, that of 'guest' access. The exact privileges of a guest user still need to be determined.

4 References

1. EGN Description of Work, can be obtained by registered users from <https://egn.bkg.bund.de/EGN.Consortium/mydms/out/out.ViewFolder.php?menuaction=dummy&folderid=20#>
2. OGC Web Feature Service Implementation Specification version 1.1.0, OGC reference code: OGC 04-094. URL: <http://www.opengeospatial.org/standards/wfs>
3. OGC Gazetteer Service application profile, OGC reference code 05-035r2, URL: <http://www.opengeospatial.org/standards/requests/36>

Appendix A: EGN service requirements

This appendix is a compilation of requirements of the EGN system from the perspectives of the different EGN stakeholders: end users, value added resellers (VARs), data providers and the EC. This compilation is to be regarded as a wish list, not as a list of things EGN is going to provide.

A.1 From the perspective of the end user

User requirements for the EGN service follow from the Description of Work and the results of WP 2 :

1. The service should be easily accessible.
2. The service should always be accessible.
3. The service should provide a fast response.
4. The service should provide understandable answers, in the language of the user.
5. It should be possible to use historical names in a query.
6. It should be possible to only query for features of a certain type (feature class, e.g. 'rivers' or 'provinces').
7. The service should be able to provide historical names in the query result.
8. The service should be able to provide variant names in the query result.
9. The service should be able to return the centroid or a single label point of each found feature.
10. The service should be able to return a geometry for each feature found (for example, a linestring, polygon or multipolygon).
11. The service should be able to return a MBR (minimal bounding rectangle or bounding box) for each feature found.
12. It should be possible to use the results from a query to make a subsequent query (iterative searches).
13. The service should provide information on the pronunciation of geographical names.
14. It should be possible to search for names that do not exactly match the name entered by the user (fuzzy searching).
15. It should be possible to search for names that sound like the name the user has entered (soundex searching).
16. It should be possible to use wildcard characters (for example "*" or "?") in the search term.
17. It should be possible to limit a search to features within a user-defined rectangle.
18. It should be possible to search for features that have certain spatial relationship (like 'overlaps', 'inside' or 'near') to a specified other feature.
19. The service should provide information on the grammatical number of geographical names.

20. The service should provide information on the gender of geographical names.

A.2 From the perspective of the VAR

It is foreseen that the primary use of the EGN service will be through Value Added Resellers (VAR). One specific example of an intermediate or added-value service that will make use of the EGN service is the ArcGIS extension that is going to be developed in WP 8. The following list of requirements was made by ESRI, who are not only going to develop the ArcGIS extension but are also involved in WP 10 (business model and marketing):

Service

1. The service should be compatible with standard GIS components.
2. The service should be of a standard service type.
3. The service should use standard interfaces for interoperability and communication.
4. Non exclusive providing (?)
5. The service should be well documented.

Hosting

6. The service should have a high availability (99,5% uptime, 24 hours a day, every day of the year)
7. The service should be fast (2,5 seconds per request).
8. The service should be monitored continuously to detect any problems.
9. Technical support should be available and respond within an hour after the report of a problem.

Business

10. EGN has to respond quickly to changes in market (e.g. prices, user-requirements).
11. EGN has to provide a fine-grained user-management to distinguish the free services from the non-free services.
12. Non-free services must have a clear benefit for the user (improved data quality or data quantity).
13. The EGN service should cover off-line- or in-house usage.

Technical

14. The EGN service should have a well-known service interface.
15. Access to the EGN service should be provided as soon as possible, so that clients can be tested.

A.3 From the perspective of the data provider

The data providers are the NMCA's. These requirements follow from communications received from the NMCA's:

1. The service should not create any additional security threats to internal data and systems.

2. The service should be usable for other purposes than the EGN gazetteer (for example: a custom gazetteer service).
3. Already existing NMCA systems, services and procedures should be used as much as possible.
4. The service should keep the 'richness' of source data intact, as much as possible.
5. The EGN software and data must be easy to maintain.

A.4 From the perspective of the EC

The requirements in this case follow from the *Description of Work*:

1. Adhere to INSPIRE directives and recommendations.
2. Develop a clear and reusable architecture.
3. Make the system extendible.
4. Use (ISO/OGC/W3C) standards where possible.
5. The service must continue to work after the EGN project has finished.
6. Single searches for geographical names within the EGN web service will be free of charge.