

15 March 2021

English

**United Nations Group of Experts on
Geographical Names
Second session
New York, 3 – 7 May 2021
Item 14 of the provisional agenda *
Toponymic data files and gazetteers**

Status report on the European Open Regional Gazetteer

Submitted by Germany**

Summary:

Since 2006, the European Union has supported projects for establishing an up-to-date data infrastructure for a regional European geographical names data service with full coverage of the European region. The Open Regional Gazetteer has been developed under the umbrella of EuroGeographics, the not-for-profit membership association for European national mapping, cadastral and land registration authorities in collaboration with its members. EuroGeographics provides European users with geographical names and other geospatial data from those authorities, as a single access point. The production management for the Open Regional Gazetteer has been taken over by the German Federal Agency for Cartography and Geodesy.

The Open Regional Gazetteer contains all geographical names extracted from two EuroGeographics products: EuroRegionalMap and EuroBoundaryMap. The data meet the requirements for geographical names of the Infrastructure for Spatial Information in Europe. Currently, the data include the geographical names of 36 European countries and “dependent territories” in 39 languages.

In addition, for those languages, standardized exonyms and variant names are linked with their corresponding endonyms and are classified according to their status.

The geographical categories covered by the Open Regional Gazetteer include a wide range of feature types, such as administrative units, hydrographic features, populated places and protected sites.

The Open Regional Gazetteer is published as a web feature service and can be obtained in GeoPackage format on request.

In addition to providing background information on the content, quality and accessibility of the Open Regional Gazetteer, the report also describes the strategical direction of EuroGeographics for that data set, including its increasing use. Through its new project, Open Maps for Europe, cofinanced by the Connecting Europe Facility of the European Union, EuroGeographics will provide direction and easy access to pan-European data sets created using official map, geospatial and land information, including geographical names and the Open Regional Gazetteer.

* GEGN.2/2021/1

** Prepared by Prepared by Jörgen Spradau and Pier-Giorgio Zaccheddu (Germany), Angela Baker (EuroGeographics), Roman Stani-Fertl (Austria), submitted by Germany

1. Background information

There has been an increasing demand for an open gazetteer service providing authoritative geographical names (GN). This demand has especially been posed by the European Commission.

EuroGeographics' ambition is to provide appropriate services with authoritative data from its members (National Mapping, Cadastre and Land Registry Authorities – NMCAs in Europe) to meet this demand. Geographical names are one fundamental geospatial data theme to serve European users. The work started under the umbrella of EuroGeographics with the EuroGeoNames (EGN) service 2006-2009 and has been significantly improved since then.

In the European Location Framework project (2013-2016) the main objective was to assemble the national INSPIRE GN services using a cascaded architecture to enable access from a single point. This work identified some issues. A cascading service faces shortcomings with performance and availability of national services. Further, national services are published with a variety of different license conditions, which are difficult to combine. The main issue is the coverage, as it gets difficult to assemble an GN service for the whole extent of Europe in short time.

In order to create a pan-European gazetteer resolving the mentioned issues, EuroGeographics has decided to reuse the toponyms (GN content) of existing EuroGeographics datasets with comprehensive European coverage. Those datasets are EuroBoundaryMap (EBM) and EuroRegionalMap (ERM), which serve as basis for the Open Regional Gazetteer.

The Open Regional Gazetteer was published in April 2019. It was made available as a web feature service (WFS) through the interface developed as part of the same project¹. Upon request it was also made available as a GeoPackage. At the end of the Open ELS project the coverage was 36 European countries and territories.

More information on the products used are available through: <https://eurogeographics.org/maps-for-europe/>

2. EuroGeographics' Strategy

EuroGeographics represents its members' interests, maintaining networks that help members improve their capabilities and role, and facilitate access to and use of members' geospatial data and services. By providing a single point of contact, Eurogeographics enables government, business and citizens to benefit from their collective expertise, datasets and services.

The organisation has traditionally has three main pillars of activity; Membership, Representation and Data.

A completely revised Data Access and Integration strategy and programme of work was developed, which put the member's requirements at it's heart. This strategy includes:

- evolution to open data integration
- emphasis on Data Access / Facilitation of members data
- building on current data offering to deliver to end users
- support for Open Data Public Sector Information (PSI) Directive

As part of its revised Data Access and Integration Strategy, EuroGeographics responded to a 2019 CEF Telecom Public Open Data call. This application was successful, and in January 2020 EuroGeographics started work on its Open Maps for Europe project². The project is coordinated by

¹ <https://www.euro-geo-opendata.eu/>

² <https://eurogeographics.org/open-maps-for-europe/>

EuroGeographics, in close partnership with the National Geographic Institute (NGI) Belgium. It also has the support of its membership, notably but not only BKG in Germany and IGN France. The project runs for three years ending in December 2022. Open Maps for Europe builds on the results of a previous project and on the experience of members. The aim is to provide easy access to free to use harmonised maps for Europe from more than 40 countries. The project supports the benefits of the Open Data Public Sector Information (PSI) Directive by driving innovation, market development and growth to support the digital economy. It also demonstrates how National Mapping, Cadastral and Land Registration Authorities contribute to the public good through a unique international collaboration.

The Open Regional Gazetteer will be continually be improved during this project. In particular, this includes:

- Improving the coverage. The coverage will be gradually improved. To achieve this, EuroGeographics will engage with its members to gain the required legal permissions to include their data within the Gazetteer and any additional country will be included within this dataset. This also includes moving the licensing permission provided by members from the ELS Data provider Agreement to the standard Framework Licensing Agreement. This will align the gazetteer to other EuroGeographics datasets.
- Increasing the use of the Gazetteer. This will be achieved by embedding it into the enhanced interface developed (to be developed as part of the project) to provide a user-friendly search function, aiding the discovery across all datasets available through the interface, as well as offering it as service through the enhanced interface.

All datasets include in this project, including the Open Regional Gazetteer will be integrated with the European data portal, enabling them to be discoverable and available. In this regard, for all open datasets, where not already created, metadata will be produced to the D-CAT (Data Catalogue Vocabulary) standards.

Key to the success of this project, is not just the technical delivery of the datasets through a new interface providing the accessibility or the opening up of the datasets through legal permission, but promoting what is available, to increase the use and re-use of this data. See Section 4 for more information on Dissemination. EuroGeographics is confident this project will be successful and the long-term view is to continue to improve the Gazetteer and promote its availability to end users. By increasing the availability and use of the Gazetteer, and other datasets, through the Open maps for Europe project, there will be a greater understanding of end user needs, which will feed into future planning. This project forms part of a wider strategy for EuroGeographics, not only delivering on our Data Access and Integration work plan, but also feeding into future work.

3. Open Regional Gazetteer - Content and structure

The following information was extracted from the Open Regional Gazetteer specification (https://thinkwhere-public.s3-eu-west-1.amazonaws.com/eurogeographics/User+Documentation/RegGaz/RegionalGazetteer_specification_1_1.pdf).

The Open Regional Gazetteer has the following properties:

- Gazetteer compliant with the INSPIRE GN specification,
- European coverage (pan-European),
- Regional level of detail (LoD) (medium scale),
- Based on the existing EuroGeographics products,
- Open data.

The Open Regional Gazetteer provides authoritative geographical names as maintained in the source data of the NMCAs. It can be used for any reference or information purpose. But, it does not imply that all data is legally approved. National specifics are described in the metadata.

The Open Regional Gazetteer is published as web feature service (WFS) on the OpenELS website. Additionally, the gazetteer is published as database in GeoPackage format. This can be requested on demand.

3.1. Source data

The Open Regional Gazetteer is derived from the EuroGeographics datasets EuroBoundaryMap (EBM) and EuroRegionalMap (ERM). EBM/ERM contain a huge amount of toponyms for a wide range of topographic feature types with full European coverage (ERM: up to 1M geographical names in 34 feature types).

The thematic completeness is sufficient for the gazetteer, e.g. EBM contains all names of administrative units, ERM contains all names of settlements (only city districts missing). The positional accuracy is absolutely sufficient for a gazetteer. ERM/EBM contain all endonyms.

All feature types with geographical names are transformed into the INSPIRE GN data model. Only reference points are used as geometry. It means, all EBM/ERM features with line or area geometry are generalised to points.

The following example visualises the transformation from EBM AdministrativeUnit to the Open Regional Gazetteer:

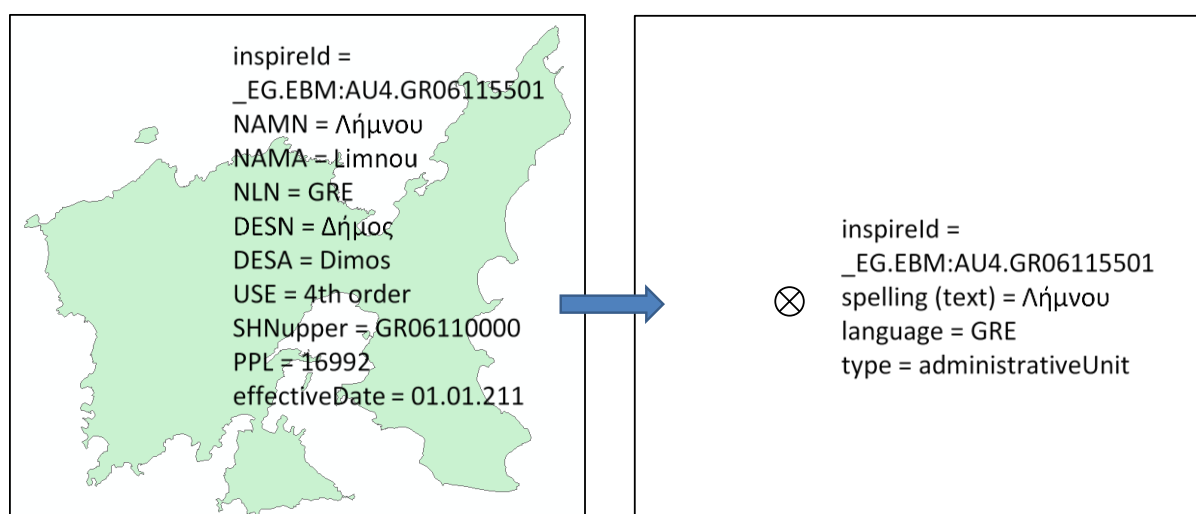


Figure 1: Transformation of an administrative unit from EBM into a named place point in the Open Regional Gazetteer

3.2. Coverage

The potential European coverage of the gazetteer is identical with the coverage of EBM and ERM, as content is derived from these existing datasets.

As the gazetteer is published as an open data service, all providers of data to EBM and ERM were asked to sign the Data Provider Agreement. If this permission is not available, the countries are temporarily excluded from the gazetteer.

The current coverage comprises the following 29 countries and territories:

Austria, Belgium, Croatia, Cyprus, Czech Republic, Denmark (including Greenland and Faroe Islands), Estonia, Finland, France (including French overseas territories and Monaco), Georgia,

Germany, Great Britain, Greece, Hungary, Iceland, Ireland, Kosovo³, Lithuania, Luxembourg, The Netherlands, Northern Ireland, Norway, Poland, Portugal, Slovakia, Slovenia, Spain (including Andorra and Gibraltar), Sweden, Switzerland (including Liechtenstein),



Figure 2: Current Coverage of the Open Regional Gazetteer

Through work already started on the Open Maps for Europe project there is an indication that this coverage will increase by at least five additional countries.

3.3. Exonyms

The project EuroGeoNames (EGN) has created a comprehensive database of exonyms and variant names, which was further extended in the European Location Framework. This initial dataset has been updated and adapted for the EBM/ERM toponyms used for the gazetteer. The focus of exonyms is on toponyms of features with major topographic relevance.

The exonyms are classified according to INSPIRE into four types (official, standardised, other, and historical), at which historical names are not yet included and only well-established exonyms in common use in the respective language are part of the database.

³ (the name Kosovo is used without prejudice to positions on status and is in line with UNSC 1244 and the ICJ Opinion on the Kosovo declaration of independence)

Example

- Endonym (Italian): Roma
- Exonym (English, French, Dutch, ...): Rome
- Exonym (German, Danish, ...): Rom
- Exonym (Welsh): Rhufain

The database comprises exonyms of an increasing number of languages. At the moment there are names in 45 languages recorded including those in minority languages like Welsh, Basque or Romansh.

A web interface has been developed for internal use to enter and edit exonyms step by step into the Open Regional Gazetteer database. The linkage to features in the gazetteer is ensured by unique identifiers. For this web interface all themes from EBM and ERM are first written in a MSSQL-Server database. After the inclusion of the exonyms, the MSSQL-Server database is transformed in an INSPIRE conform PostgreSQL/PostGIS database.

The focus of exonyms is on toponyms of features with major topographic relevance. At starting point for exonyms and variant names, the selection criteria are given in the following table:

EBM/ERM feature type	Selection criteria for toponyms
BuiltupP	<ul style="list-style-type: none"> • Settlements with more than 100000 inhabitants, • Capitals and residences of authorities of Administrative Units (2nd, and 3rd order).
AdministrativeUnit_1	<ul style="list-style-type: none"> • Country names in their short form
AirfieldP/C	<ul style="list-style-type: none"> • English name forms

The linkage of exonyms and variant names with the respective endonyms will be extended step by step. The density of exonyms entered varies by language. It is depending on the availability and quality of list of exonyms in the respective language provided.

According to the extracted data from ERM / EBM the database holds toponyms in four different scripts (Latin, Cyrillic, Greek, Georgian). Exonyms in Non-Roman scripts have two entries. They are entered in their original spelling and they are Romanised by the conversion system recommended by UNGEGN. If such a system does not exist, the national recommended system is used.

3.4 Data Model

The Open Regional Gazetteer utilizes the INSPIRE data model GN. INSPIRE GN aims at providing a geographical names dataset as a gazetteer comprising geographical names specific attributes if available. The UML diagram is visualised in figure 3.

This document does not include a feature catalogue – please refer to the INSPIRE Data Specification on Geographical Names (<https://inspire.ec.europa.eu/id/document/tg/gn>).

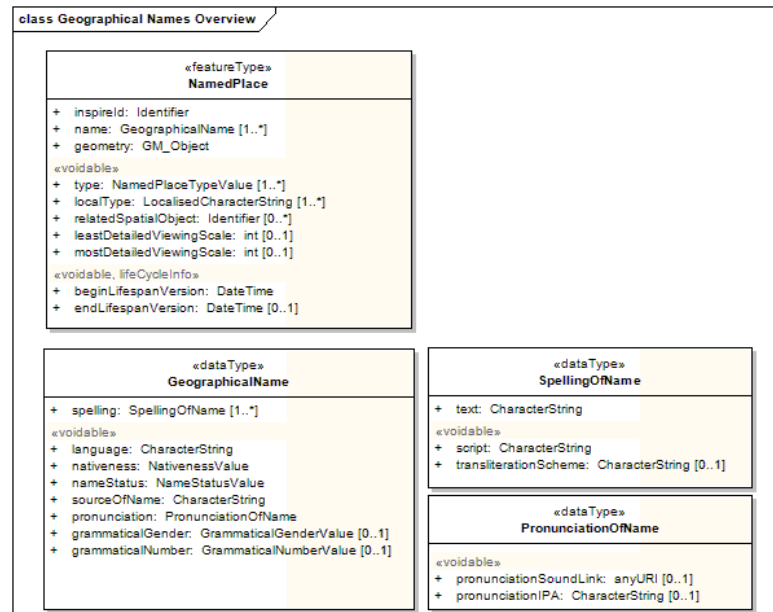


Figure 3: INSPIRE data model GN

The basic feature type of the INSPIRE data model GN is *NamedPlace*, which stores the geometry. Each *NamedPlace* has one or more *GeographicalName* (either endonyms or exonyms). Each *GeographicalName* has one or more *SpellingOfName* with different script (Latin, Greek, ...).

Example of the Greek capital Athens (only 2 of 37 exonyms listed):

- NamedPlace
- type = populatedPlace
- localType = Built-up Area / Populated Place
- GeographicalName (1)
 - language = gre (Greek)
 - nativeness = endonym
 - SpellingOfName (1)
 - text = Αθήνα
 - script = Grek (Greek)
 - SpellingOfName (2)
 - text = Athína
 - script = Latn (Latin)
 - transliterationSchema = ELOT 743
- GeographicalName (2)
 - language = eng (English)
 - nativeness = exonym
 - SpellingOfName
 - text = Athens
 - script = Latn (Latin)
- GeographicalName (3)
 - language = geo (Georgian)
 - nativeness = exonym
 - SpellingOfName
 - text = ათენი
 - script = Geor (Georgian)

The attribute *type* of *NamedPlace* is populated in the following way:

NamedPlace localType (EBM/ERM)	NamedPlace type
Built-up Area / Populated Place (BuiltupP), Named Location (UrbanP)	populatedPlace
Airport / Airfield (AirfldP/C), Ferry Station (FerryC), Port (HarborP/C), Railway Station (RailrdC), Entrance / Exit (ExitC)	transportNetwork
AdministrativeUnit_x and designation	administrativeUnit
Lake / Pond / Reservoir (LakeresA), Island (IslandA), Watercourse (WatercrsL), Glacier / Snow Field / Ice Field (LandIceA)	hydrography
National Park / Nature Reserve (ParkA)	protectedSite

The attribute *relatedSpatialObject* of feature type *NamedPlace* is populated for watercourses, where it is of major interest to select all named places and geographical names which refer to one real world object. This linkage method is used for all cross-border features.

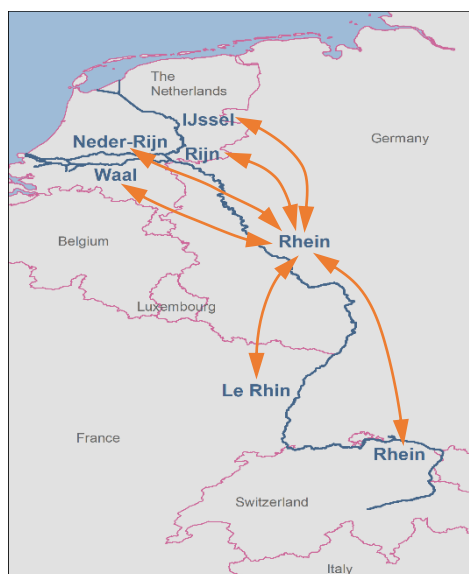


Figure 4: Example of implementation for the Rhine River

3.5 Data quality

In general, data quality elements are described in the EBM and ERM specifications.

It should be considered that the Open Regional Gazetteer will contain some name duplications due to multiple name occurrences and to multiple entries of the same object/name but with different ids.

Metadata are published on the European Data Portal (<https://www.europeandataportal.eu>).

4. Dissemination and Awareness

To be successful in the Open Maps for Europe project and beyond, EuroGeographics must raise the awareness of what datasets are available and what they can be used for, resulting in increased use and an extension of the end user base. Through the Open Maps for Europe project, the new interface and open datasets will be promoted, paying particular attention to demonstrating the ease of access to

the datasets, and how they can bring value to any business or citizen.

By means of social media (including Twitter, LinkedIn and Facebook), press releases, animations/videos and attendance at events, EuroGeographics will promote the interface and its datasets to end users and stakeholders increasing the take-up over the duration of the project. In addition, EuroGeographics will engage with its members to request the open datasets are linked to their national geoportals, which will increase the visibility and maximise accessibility, particularly within their national bodies and ministries.

Case studies will also be developed, published and promoted for key sectors as part of the project. These case studies will demonstrate the value of the reuse of open geospatial public sector information, and the aim is to include one on the Open Regional Gazetteer.

EuroGeographics has clear evidence from its open dataset, EuroGlobalMap, that once datasets have been made available as open data, and are straightforward to access, the use significantly increases. EuroGeographics is confident this will happen with all datasets in the Open Maps for Europe project, including the Open Regional Gazetteer.

4.1 Licensing

The dataset and service are currently licenced under the Open licence developed as part of the ELS project: <https://www.euro-geo-opendata.eu/els-licence>. As part of the Open Maps for Europe project, the open regional gazetteer will be licensed through the standard EuroGeographics open data license, ensuring that all EuroGeographics open datasets are following the same license. This is to make it more straightforward for EuroGeographics members and the end users.

The attribution guide is published on the Open user interface: <https://www.euro-geo-opendata.eu/service/open-regional-gazetteer-service>

Although this dataset is utilising data from existing EuroGeographic's Pan European datasets, it is published as an Open service, and therefore the project requires that data providers include permissions for this data to be used for this dataset within the Data Provider Agreement.

The gazetteer will be updated according to the update cycle of EBM and ERM. These source datasets are updated annually.

For enquiries contact: angela.baker@eurogeographics.org

5. Outlook

It is clear from feedback from the European Commission that a pan European Gazetteer is a requirement. EuroGeographics, with the support of its members, is in an ideal position to offer this. With the Open Maps for Europe project, and the opening up of further datasets, the Gazetteer will be made easily accessible and be heavily promoted. This will increase use, but also user feedback. This feedback will enable further development planning to be done, with the aim of continually improving the Gazetteer.

Points for discussion

The Group of Experts is invited to:

- (a) Take note of the report and progress made by EuroGeographics towards a Regional Gazetteer for Europe
- (b) Express its views on the way forward concerning concept, the development as well as extension and development on the Open Regional Gazetteer