UNITED NATIONS
GROUP OF EXPERTS
ON GEOGRAPHICAL NAMES

WORKING PAPER
NO. 22/9

Twenty-eight session
28 April – 2 May 2014

Item 9 of the Provisional Agenda

Activities relating to the Working Group on Toponymic Data Files and Gazetteers

Status Report on the EuroGeoNames (EGN) Project – Implementing a sustainable services infrastructure within the European Location Framework (E.L.F.)*

* Prepared by Dr. Saulius Urbanas (Lithuania) EuroGeoNames Project Manager, submitted by Pier-Giorgio Zaccheddu (Germany), Convenor of the Working Group on Toponymic Data Files and Gazetteers of UNGEGN

Status Report on the EuroGeoNames (EGN) Project – Implementing a sustainable services infrastructure within the European Location Framework (E.L.F.)**

Summary

EuroGeoNames (EGN) is a service providing the combined geographical name services from 17 National Mapping and Cadastral Authorities (NMCAs), members of EuroGeographics, the European Association of NMCAs. EuroGeoNames provides authoritative data and includes official exonyms and variant names representing the 25 languages of the EU.

In order to simplify the EGN architecture and making the service INSPIRE compliant the major renewal of EGN service was executed in 2011-2012. The renewed service is easier supported by data providers, more user friendly and provide additional functionality. An access to the EGN service is available for free from EuroGeographics website www.eurogeographics.org [1].

EuroGeographics continues to work with its members, European NMCAs, in extending coverage of EGN by linking national web-services of the geographical names and further working on the quality improvement, maintenance and the population of the exonyms and other variant names database associated to geographical names through EGN service. The national contributions to EGN are provided on the basis of common Framework License Agreement.

Now the EGN service directly contributes to the development of GeoLocator service in the European Location Framework (ELF) – the sustainable cloud based infrastructure - providing users and application development an access to up-to-date, authoritative, interoperable, cross-border, reference geo-information for use by the European public and private sectors.

1. Background

EuroGeoNames (EGN) started as an eContent*plus* project in 2006. Following successful completion of the project in 2009 EuroGeographics decided to invest in the further develoment of the service. Since 2009 EuroGeographics has been working to increase coverage of the service and has studied the market possibilities and user requirements via the ESDIN project. This work identified changes needed in the architecture of the service.

In 2009-2012 the EGN Central Service was hosted by EuroGeographics together with the German Federal Agency for Cartography and Geodesy (BKG), and was implemented as a Web Feature Service (WFS) node using an Open Source software product called degree. The database is based on PostgreSQL relational database and on its spatial extension called PostGIS. The national services and databases are implemented using the same tools. The update mechanism of the centralized database was based on periodic update, in which the data set related to a certain country was completely replaced by the new full data set retrieved from the national service. The dataset was imported into the centralized database using the functionalities of the transactional WFS-T interface and some further database scripting. The original EGN schema, and the database structure derived from this schema, was developed as an extension to the ISO 19112 Gazetteer schema (ISO 19112 Geographic information – Spatial referencing by geographic identifiers). As a result of the data modelling work done in the EuroGeoNames project, the EGN schema was expanded significantly, when compared with the original ISO 19112 schema. The database structure for the EGN Central Service, and for the national service nodes too, has been derived from this schema.

Over the years it has become obvious that the EGN schema is too complicated. There were a couple of facts that confirm this assessment:

- The output schema was considered too heavy for the Web applications running in a Web browser-based environment
- The database structure of the EGN data store was very difficult to understand and manage
- The data update procedures were difficult due to the overly complicated data structures; this is evidenced by the fact that data updates in the Central Service have resulted in corrupted data.
- It has been difficult to convince new countries to join EGN service, partially because of the very demanding data structures present in the current national EGN databases

2. Renewal of EuroGeoNames Central Service

After the investigation in to possibilities for renewal of EGN service following tasks were defined by EuroGeographics involving experts from NMCAs [4]:

- Technical implementation of the new EGN Central DB schema
- Development of the transformation from the current EGN schema to the new EGN Central DB schema
- Populating the new EGN Central DB from the national data sets and the exonym data set
- Development and technical implementation of the OGC Gazetteer Service AP-compliant service on the top of the new EGN Central DB schema

- Establishment of the Cloud Service-based EGN Central DB and OGC Gazetteer Service AP-compliant instance

The EGN renewal work on demand of EuroGeographics was executed by Finnish Geodetic Institute. The simplified EGN Central DB schema was implemented using PostgreSQL/PostGIS database management system. The main content of the EGN data set are stored in two tables 'si_location_instance' and 'alternative_geographic_identifier'. The table 'si_location_instance' contains all the data items that relate to geographic places (location instances in ISO 19112 terminology). The table 'alternative_geographic_identifier' stores the place name data. The linkage from the names to the corresponding place is established by creating an association between the two tables (using primary key and foreign key based connections). The tables 'si_gazetteer' and 'si_location_type' store the information related to the data set itself (gazetteer) and the location types (feature classification) supported by the gazetteer. The remaining two tables 'gazetteer_location_type' and 'location_type_link' are supporting tables that help in managing the location type information [4]. The EGN Database Schema is presented in Figure 1.

A transformation was developed from the previous EGN schema to the new OGC Gazetteer Service AP schema to help in data upload process. This transformation is based on XSLT technology and was used as an internal function of the Java-based data upload process.

The national EGN services were accessed to download all of their content. Then the downloaded data sets were transformed to the new schema and uploaded to the central database using a custom built Java-process. The exonym data set was processed separately, as it was stored in a completely different structure using different technical solutions [4].

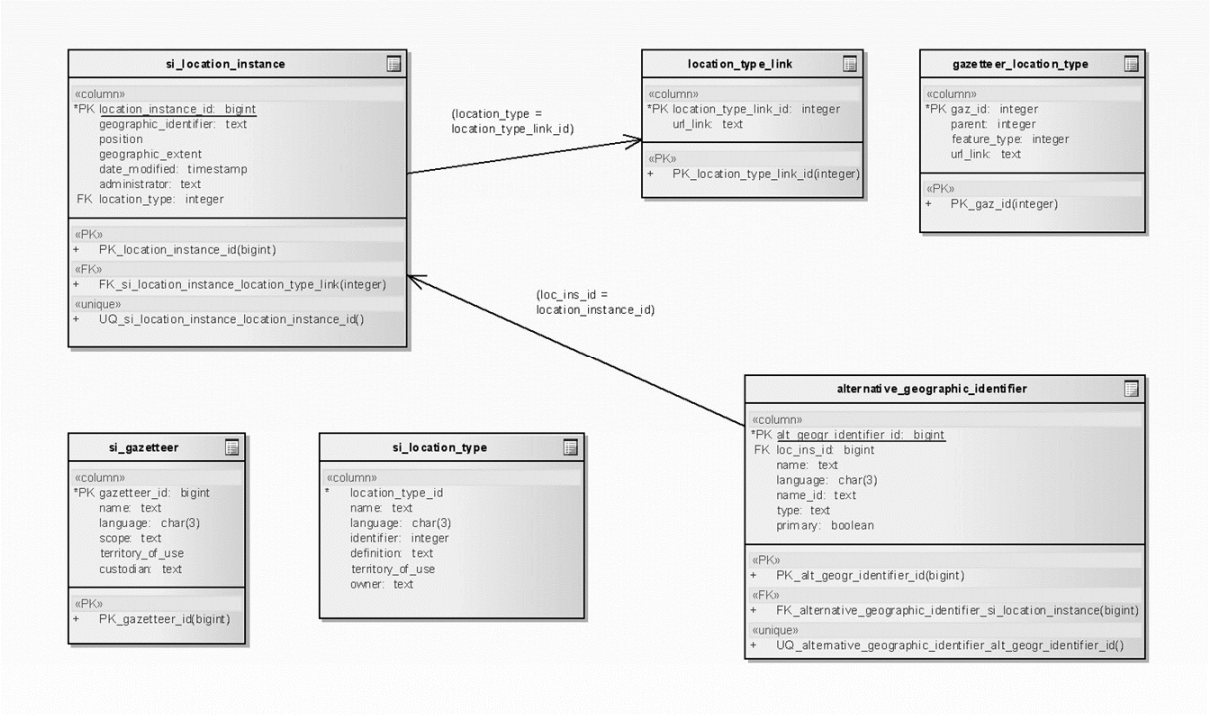

Figure 1. EGN database schema

OGC-specified Gazetteer Service was established based on deegree platform. Transformations needed for complete schema validity were implemented partly by using deegree's schema transformation capabilities, partly by a custom-built front end application.

Full schema validity was achieved in the output of the service. Extensions needed for support on multilingualism were implemented according to the principles established in the INSPIRE-process (LANGUAGE –parameter). The Gazetteer Service was established as Cloud Service-based instance in the Amazon Web Service –platform [4].

The renewed EuroGeoNames service provides new features as [1]:

- OGC standartised Web Map Tile Service (WMTS)
- access to EGN content in Java Script Object Notation (JSON) format
- support for KML-encoded output. In addition the service can be accessed directly from a network link that could be visualised in the Google Earth
- Output of EGN Central Service is available as Core Location Vocabulary (CvL) via the Web Feature Service (WFS), or the Gazetteer Service AP of WFS, interface.

3. Development of European Location Framework

The European Location Framework (ELF) is the project of the Consortium of 30 partner organisations (14 NMCAs, EuroGeographics, system integrators, application developers, user community representatives) co-funded by European Commission under the Information and Communication Technologies (ICT) Competitive and Innovation (CIP) programme. The European Location Framework (E.L.F.) project will deliver a first implementation of a technical infrastructure which enables users to gain access to authoritative, interoperable and harmonised geospatial reference data across Europe for use by the European public and private sectors. It will foster the wider use of geo-information and enable the creation of innovative value-added services. [2]

The project will provide a critical mass of content and coverage as 15 Member States' national INSPIRE data will be made available from a single point. Covering the full range of INSPIRE Annex I,II and III themes, these datasets will provide full national coverage of the rich content available from national and regional spatial data infrastructures [3].

ELF will enable users to access authoritative spatial data from the NMCAs based on common specifications through the ELF Platform. The common INSPIRE compliant ELF data specification based on extensions to existing INSPIRE rules and guidelines will be agreed and then applied via Geo-tools to existing local national services to enable cross-border interoperability [3]. The ELF platform's Geo-tools will be based on extensions to existing services (as developed in ESDIN) for the required data transformation, validation (quality evaluation), generalisation and edge-matching.

The output ELF reference geo-information datasets will be published via cloud GIS service platforms (Open Source Oskari Platform and ArcGIS Online), which will be available and maintained for valued added application development. Geolocator service will provide georeferencing services needed for the ELF platform. It will be based on geographical names, addresses and administrative units.

The ELF delivers NMCAs contribution to European Spatial Data Infrastructure and e-government which is one of the ambition behind the INSPIRE directive, supports the directive on the re-use of Public Sector Information, provides the mechanism to satisfy the Reference Data Access component of Copernicus, is an element in the European Interoperability Framework, and is an important contribution to the Digital Agenda for Europe. More information about the ELF project – www.elfproject.eu

4. Utilisation of EuroGeoNames service by ELF GeoLocator Service

The data contents of the ELF GeoLocator Service comprise of a EuroGeoNames (EGN) service data and data that has been retrieved from the ELF participants'INSPIRE/ELF Download Services that cover the themes Addresses (AD), Administrative Units (AU) and Geographical Names (GN). The EGN data includes the place names from a special exonym database that contains multilingual names from several important European locations (Figure 2). The AU and AD data are inserted to the database as new place name instances using the name of the AU or the address information as the name content [5].

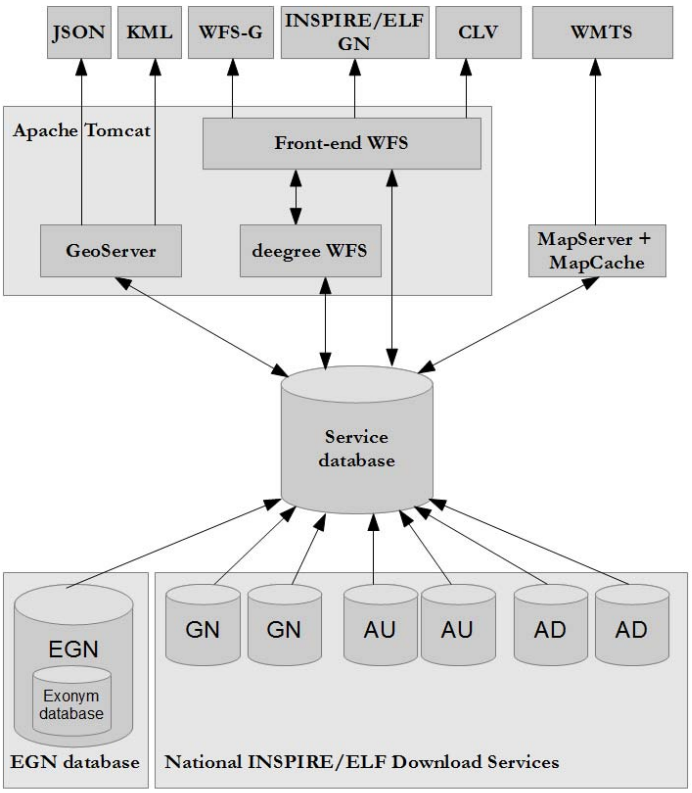

Figure 2. The architecture of GeoLocator Service

The GeoLocator Service is based on centralized architecture, in which the data sets are collected from national level sources and stored in a single service database (see Figure 2) [5]. The main access method for harvesting the national data sources is to use the INSPIRE-compliant Direct Access Download Service interface, i.e. the Web Feature Service (WFS). Various Open Source tools are used to produce the different output encodings. A custom-built front-end module takes care of certain functionalities that are not supported by the tools used. An exemplary end user application has also been developed to demonstrate the functionalities offered by the GeoLocator Service. The ELF GeoLocator Service provides functionalities for performing geocoding and reverse geocoding. Geocoding refers to the process of finding the coordinates of a location from the basis of a given place name or address. The reverse geocoding refers to the process of finding the place name that is located nearest to the given coordinate point. The geocoding functionalities of the ELF GeoLocator service include ordinary geocoding, administrative unit -limited geocoding and fuzzy name searchbased

geocoding. The reverse geocoding contains two functionalities, ordinary reverse geocoding and administrative unit-limited reverse geocoding.

The main output format of the GeoLocator Service is the GML output that is compliant with the schema defined in the Open Geospatial Consortium's (OGC) Gazetteer Service Application Profile of the Web Feature Service (WFS-G AP) Best Practice document [2]. The output of the GeoLocator Service is provided mainly as SI_LocationInstance features, defined in the Gazetteer Service AP. Other supported GML-based output forms are the INSPIRE/ELF GN output and the Core Location Vocabulary (CLV) output [6]. Other supported output formats are JavaScript Object Notation (JSON) and Keyhole Markup Language (KML). The place name contents of the GeoLocator Service are also provided in visual form as a Web Map Tile Service (WMTS) [7] (Figure 2).

References

[1] http://www.eurogeographics.org/eurogeonames

[2] www.elfportal.eu

[3] European Location Framework – One Reference Geo-Information Service for Europe, A. Jakobsson, ICC Dresden, 2013

[4] Renewal of the EGN Central Service, Finnish Geodetic Institute, 2012.

[5] ELF GeoLocator Service, draft deliverable D3.4, ELF project, 2014.

[6] Gazetteer Service – Application Profile of the Web Feature Service Best Practice / version 1.0 / 30 Jan, 2012 / Harrison, J., Vretanos, P. A. (ed).
https://portal.opengeospatial.org/files/?artifact_id=46964

[7] OpenGI S Web Map Tile Service Implement at on Standard/1.0.0/6 April, 2010/ Maso, J., Pomaki s. K. and Nśria, J.
http: //portal.opengeospatial.org/files/?artifact_i d=35326