23 June 2017

Original: English

Eleventh United Nations Conference on the
Standardization of Geographical Names
New York, 8-17 August 2017
Item 12 (b, c) of the provisional agenda*
Toponymic data files and gazetteers:
    b)   Data management and interoperability;
    c)   Data services, applications and products (for example, gazetteers and web services).

Maintaining and publishing the geographical names database (GN-DE) within the German Spatial Data Infrastructure (GDI-DE)

Submitted by Germany**

Introduction

In Germany, all geographical names information of several (topographic) vector data products based on scale level 1:250,000 and smaller is maintained in a single database called 'Geographical Names of Germany (GN-DE)'. Through unique identifiers the names entries of the GN-DE are linked unambiguously to all databases and products provided by the Federal Agency for Cartography and Geodesy (BKG). The administration and maintenance tasks of the GN-DE are conducted in Frankfurt am Main at the headquarters of BKG, the German national mapping agency responsible for small to medium scale mapping, whereas the publication of the geographical names data is facilitated through a web (gazetteer-) service by the BKG Central Service and Distribution Center for Geoinformation in Leipzig [1].

Today, GN-DE includes all relevant object/feature types from the digital landscape model scale 1:250,000 (DLM250), which have one or more geographical names (for reasons such as multilingualism).

1. Maintaining and publishing the national geographical names database (GN-DE)

The object/feature themes of the GN-DE geographical names database are: ‚settlement‘, ‚transportation‘, ‚vegetation‘, ‚hydrography‘, relief‘, areas and regions‘ with attributes regarding the administrative hierarchy (administrative code), river system (hydrographic key number in accordance with the German Working Group on water issues (Bund/Länder-Arbeitsgemeinschaft Wasser - LAWA), number of inhabitants, size (area, length), position (geometry data from DLM250 as well as additional attributive geographical coordinates), postcodes, status, language of the geographical name and height data (in meters).

The data model GN-DE was developed and has been evolved continuously by BKG following UNGEGN resolutions. Today, all geographical names entries of the GN-DE are linked unambiguously to all databases and products provided by BKG through unique identifiers (UID).

GNObjekt

nnid	landescode	beschreibung	geolaenge	geobreite	hoehe	groesse	ewz
DEBKGGND00000IUD	276		14,4276880830704	51,1808680587142		67	39845

Endonym

nnid	name	geschlecht	sprache_id	status_id
DEBKGGND00000IUD	Bautzen		1	1
DEBKGGND00000IUD	Budyšin		13	1

Status 1 = official

Language 1 = german
Language 13 = upper sorbian

DLMLink

ui_id	nnid
DEBKGDL100000EAQ	DEBKGGND00000IUD
DEBKGDL200000WJK	DEBKGGND00000IUD
DEBKGDL2000017O7	DEBKGGND00000IUD

PLZ

nnid	plz
DEBKGGND00000IUD	02625

Figure 1 - GN-DE database table showing that one GNObject may be associated to one or more geographical names (endonyms) in different languages. The example shows the GNObject associated with 'Bautzen' (in German language) and 'Budyšin' (in Sorbian languages)

At present the GN-DE database contains about 160,000 entries including more than 1300 names in the languages of the national Sorbian (Upper and Lower), Frisian and Danish linguistic minorities. About 27,000 entries indicating the gender of hydrographic features are contained as well. BKG has been providing geographical names to the general public through the Internet since 2006. The gazetteer web service application 'Geographical names' is based on a Web Feature Service (WFS) called 'WFS GN-DE' and is thus compliant with the corresponding Open Geospatial Consortium (OGC) specification [2].

With the web service application you can search the geographical names contained in the dataset GN-DE online.

Find a Name

Berlin [Find]

Name (min.2 characters
* for arbitrary characters)

Map Presentation and Search

Name	Object Type	LAT	LON
Berlin	AX_Ortslage	13°24'20"	52°31'4"
Berlin		10°26'52"	54°2'10"
Berlin	AX_Gemeinde	13°26'47"	52°31'4"
Berlin	AX_Bundesland	13°26'47"	52°31'4"
Berlin	AX_KreisRegion	13°26'47"	52°31'4"
Berlin	AX_Verwaltungsgemeinschaft_ATKIS	13°23'21"	52°30'39"

Height	Height_calc	Area	Population	Language	Status	Map
35 m	34 m	892 km²	3421829	deutsch	amtlich	
	37 m		474	deutsch	amtlich	
			3421829	deutsch	amtlich	
			3421829	deutsch	amtlich	
			3421829	deutsch	amtlich	
			3421829	deutsch	amtlich	

Figure 2 – Screenshots from the German geographical names application based on a Gazetteer Web Feature Service (WFS)

2. GN-DE within the German Spatial Data Infrastructure (GDI-DE)

The GN-DE dataset is currently used for the geographical names data provision within the '*German Spatial Data Infrastructure* (GDI-DE)'. GN-DE is the rationale for the INSPIRE data provision for European purposes as it is the only dataset in Germany that includes the mandatory and almost all 'voidable' INSPIRE attributes (like language, status, gender …). A wide range of different applications (e.g. geoportals) has been built up that provides discovery functionality for spatial data sets and services and generally make use of the functionality. In Germany, the Geoportal.de provides insight into GDI-DE. It utilizes several Web Services from GDI-DE. Users can promptly search within the central search engine of GDI-DE, which currently contains around 87,000 decentrally maintained sets of metadata on spatial data and spatial data services from across all levels of public administration in Germany. The Geoportal.de application is completed by a search utility for places and addresses.

This search utility has been developed at the BKG Central Service and Distribution Center for Geoinformation and is widely used in the portal application. It builds upon the '*Georeferenced Address Data Federation*' data set and parts of the very comprehensive names dataset 'Geographical Names of Germany' (GN-DE) derived as dataset 'GN250', Geographical Names of Germany (1:250,000) [3]. As mentioned before, GN-DE provides the unique source for all BKG products and naturally for the national (gazetteer) web feature service, too. GN-DE also provides geographical names in the officially recognized Sorbian and Frisian minority languages. The integration of those

minority language names in the search utility for places and addresses of the Geoportal.de is envisaged, but not realized as yet (2016).
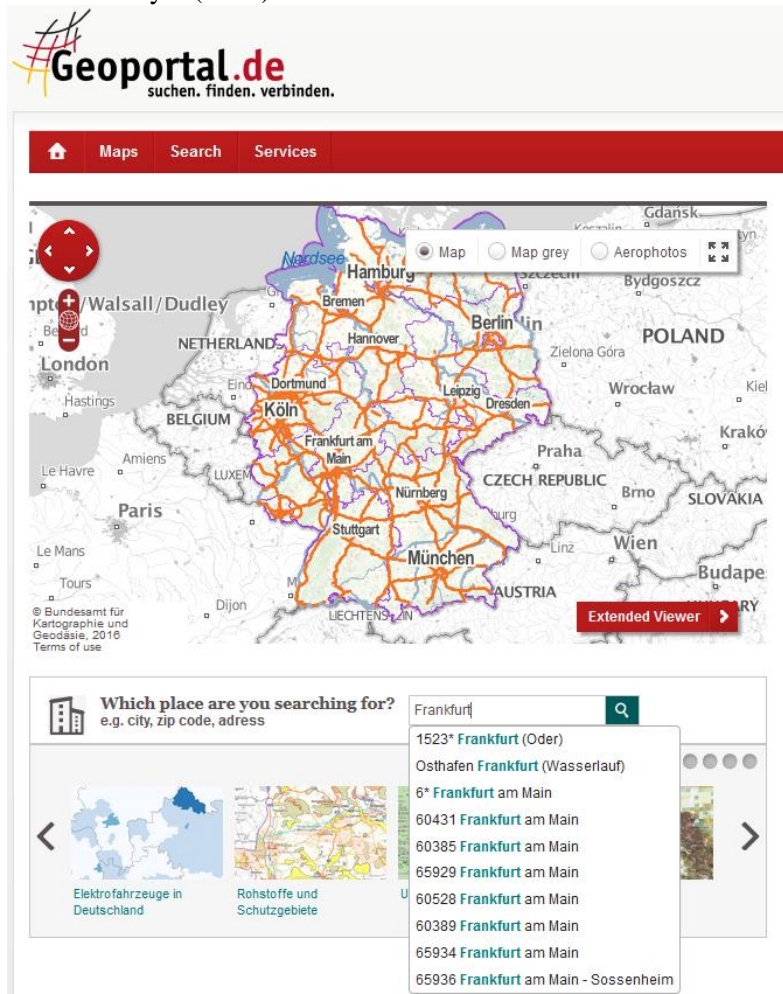

Figure 3 - The German national geographical names database (GN-DE) published as a web service which is part of the national spatial data infrastructure (GDI-DE) and visualised through the Geoportal application (Geoportal.de)

The 'Georeferenced Address Data Federation' dataset mainly consists of data from the 'Association for the Distribution of House Coordinates' run by the surveying authorities of the Länder. After processing, this dataset provides point coordinates for house addresses and encircling bounding boxes of streets, places and zip codes. The search utility implements two search strategies: a high performance live suggest search, i.e. to deliver real time suggestions for search terms and/or other relevant information based on live user input in an application, on the one hand, and a fault-tolerant search on the other. Words matching the search term are highlighted. While in the results of the live suggest search only the elements of the objects' name as part of the address are shown, the fault-tolerant search provides the complete dataset including geometry. If the match between search term(s) and results is weak, a list of similar places or addresses is offered. The search utility is able to handle unstructured input of addresses.

Figure 4 - The search utility within the Geoportal.DE – live suggest during user input

The search utility has been implemented as Java-Servlet and offers an HTTP/Key-Value Pair interface. For data indexing the Apache Lucene library is used [4]. For geometry handling and implementation of OGC-standards the GeoTools library is used [5]. Overall, there is a total of 23 million indexed datasets with an aggregate data volume of approximately 3 GB. The index is built once before service start-up and is loaded into the central memory.

Meaningful combinations of map layers available within GDI-DE are placed prominently on the Geoportal.de Welcome page as featured maps. Especially by overlaying spatial information layers from different institutions the value added through the use of interoperable spatial data services is demonstrated. The selection of thematic maps covering Germany will be extended continually.
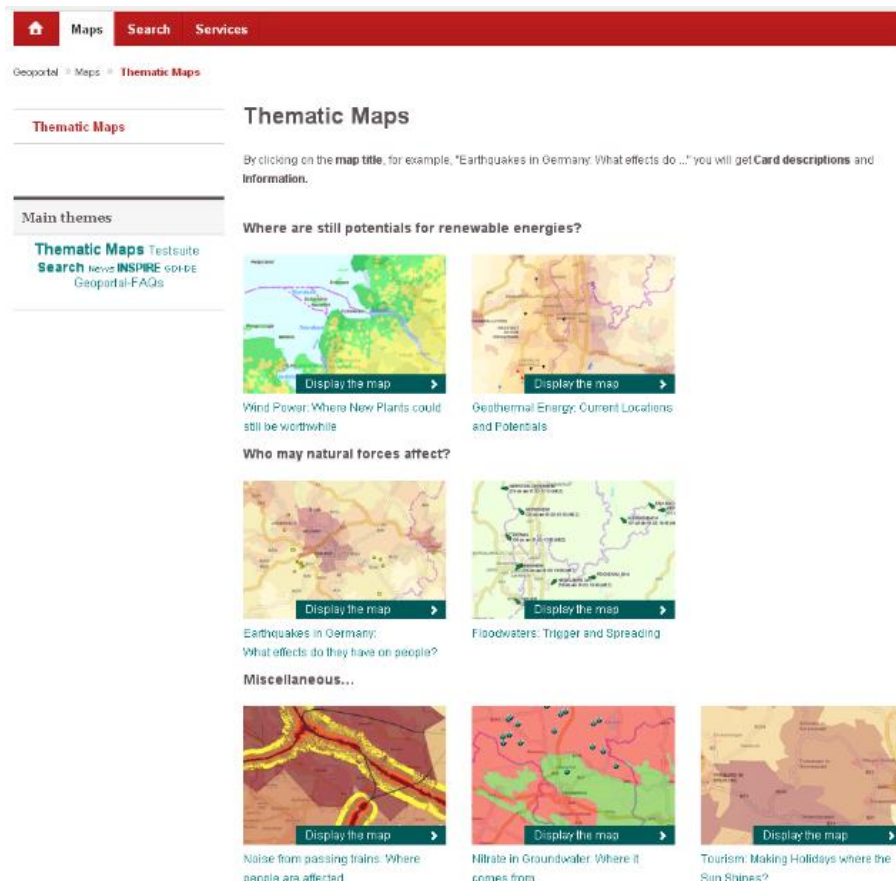

Figure 5 - Geoportal.DE as a showcase for Spatial Data Infrastructure (SDI) through featured maps

Recently the physiogeographic regions of Germany combined with touristic information were integrated [6]. The development of this service is described in another Conference Paper. In this way, one important product created by the German Permanent Committee on Geographical Names (StAGN) will be published and disseminated in the broader GDI-DE context. Considering multilingualism and the cultural heritage in Europe, people often use different spellings and languages when talking about one and the same place. Even within a country more than one (official) spelling is used. These INSPIRE-supported principles also underline the United Nations resolutions on the standardization of geographical names which are reflected in Germany in the work of the Permanent Committee on Geographical Names (StAGN).

3. GN-DE brought into conformity to INSPIRE

In 2015, the geographical names database GN-DE and the Implementing Rule for ‚Interoperability of Spatial Data Sets and Services' (this is, data specification of geographical names, Annex I) of the European geospatial data infrastructure (keyword: INSPIRE) have been brought into conformity [7]. According to the INSPIRE schedule EU countries are only obliged to realize this by the year 2017.

INSPIRE GN data schema:

Figure 6 - INSPIRE Data schema fulfilled by BKG

The INSPIRE-conform geographical names dataset of BKG is already available via the European data portal (http://www.europeandataportal.eu/de/). The European data portal collects metadata from information of the public sector which is available in European countries. This also includes information regarding the provision of data and the advantages of their further use.

4. GN-DE published as Open Data

Every day, public authorities all over the world gather a multitude of data and information such as spatial data on transport networks, addresses and cadastral information or economic, environmental and statistical data. In the recent past more and more government authorities admit to the principle of "Open Government" and "Open Data" policy.

The Geodata Access Act (Geodatenzugangsgesetz, GeoZG) passed in Germany by the federal government on 30 July 2008 promotes the further development of the German Spatial Data Infrastructure (GDI-DE). Germany declared the federal government's geodata to be open data in the course of the revision to the Geodata Access Act (GeoZGÄndG) on 16 November 2012. The Ordinance to Determine the Conditions for Use for the Provision of Spatial Data of the Federation (Verordnung zur Festlegung der Nutzungsbestimmungen für die Bereitstellung von Geodaten des Bundes) (GeoNutzV), which allows commercial and non-commercial use of the data free-of-charge, was passed on 19 March 2013. In particular, the GeoNutzV ordinance lays down provisions relating to the provision of spatial data, and to spatial data and metadata services on behalf of spatial data holding agencies [8].

In Germany, the open data declaration refers to geospatial data produced by the Federal Agency for Cartography and Geodesy (BKG). Consequently, view and download services based on international open standards and open source software of digital landscape model (DLM) and digital terrain model (DGM) as well as digital topographical maps (DTK), general maps and maps of municipality areas based on small to medium scale are provided free of charge. Hence, the gazetteer web service of BKG is available free of charge, too.

5. References

[1] Central Service and Distribution Center for Geoinformation at BKG, www.geodatenzentrum.de, last accessed 04/2017
[2] Geographische Namen Deutschland – GN-DE, the (gazetteer) WFS service (wfs_gnde), http://www.geodatenzentrum.de/xml/WFS_gnde.xml, last accessed 04/2017
[3] Geographische Namen 1:250,000 (GN250), documentation of the dataset (in German only), http://www.geodatenzentrum.de/docpdf/gn250.pdf, last accessed 04/2017
[4] Apache Lucene: http://lucene.apache.org, last accessed 04/2017
[5] http://geotools.org/ , last accessed 04/2017

[6] Bundesrepublik Deutschland 1:1,000,000, Landschaften - Namen und Abgrenzungen, 6st edition (Federal Republic of Germany at 1:1,000,000 scale, Geographical regions - names and boundaries, 6st edition. Federal Agency for Cartography and Geodesy (BKG)

[7] INSPIRE Implementing Directive (COMMISSION REGULATION - EU - No 1089/2010 of 23 November 2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services), http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:323:0011:0102:DE:PDF, last accessed 04/2017

[8] Ordinance to Determine the Conditions for Use for the Provision of Spatial Data of the Federation (Verordnung zur Festlegung der Nutzungsbestimmungen für die Bereitstellung von Geodaten des Bundes – GeoNutzV): http://www.geodatenzentrum.de/docpdf/geonutzv_eng.pdf, last accessed 04/2017