

Section 2 Management of a national names programme:

Chapter 3 Maintaining and publishing the geographical names database in Germany

Pier-Giorgio Zaccheddu

3.1 Introduction

In Germany, all geographical names information of several (topographic) vector data products based on scale level 1: 200,000 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). At present, the database contains about 165 000 entries including more than 700 names in the minority languages Sorbian and Frisian. 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 Centre for Geoinformation in Leipzig [1]

The following information summarizes the tasks and resources needed to set up and maintain the management of a national names programme. The figures are based on rough estimations as the logistics, production time needed or necessary upkeep costs can vary significantly according to the time and resources

needed for the adaptation of the database, the services and applications to new technical developments or to the services-oriented architecture in the organization.

3.2 The Geographical names standardization process in Germany

In Germany, a decentralized responsibility between the federal government and the Länder (federal states) exists and this must be considered for the elaboration and assessment of the management of a national geographical names programme. The government in Germany has three distinct levels of public authority: local, regional and national, all of which are generators and holders of public and geospatial information. Surveying and mapping is the task of the 16 Länder. Each of the Länder is responsible for its own topographic and cadastral service, environmental and statistical data collection, and in general for data policies. Traditionally, data collection is largely decentralized and carried out mostly on the regional and local level, which means that the processing and maintenance of data is mostly tailored to local and regional requirements, which leads to a built-in incompatibility. The different Länder have issued laws that regulate the work, which the regional and local authorities are carrying out. The "Surveying and Cadastral Acts" may serve as an example. The surveying and mapping administrations of the Länder are responsible for creating and maintaining the geospatial (reference) data describing real estate and the landscape, including geographical names. Whereas the mapping agencies of the Länder are responsible for providing large and medium scale reference data, the cadastre offices in the Länder support them by performing the tasks of the real estate cadastre.

The responsible regional mapping authorities are collaborating in the Working Committee of the Surveying Authorities of the Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland – Adv [2]). The Adv provides joint documentation regarding the geospatial reference data available as well as a joint pricing and data collection policy. ATKIS® (Authoritative Topographic-Cartographic Information System), the central topographic programme of the German national surveying authorities and BKG. The product line comprises digital landscape models (DLM), digital terrain models (DGM) and the digital topographic map series 1: 25,000 to 1: 1,000,000. There is an authorization by agreements between the Federal Administration and the Länder on the production of topographic reference data (including maps). Amongst others, all geospatial reference data sets based on scales larger than 1: 200,000 are done by the federal states, while the data sets with equal and smaller scales are compiled by BKG.

The 'Bundesgeoreferenzdatengesetz (BGeoRG)', the Federal Geospatial Reference Data Act, regulates the quality proofed and standardized provision of geodetic reference systems and frameworks as well as of the geospatial reference data products of the federal administration. Currently, geospatial reference data has varying degrees of quality e.g. in terms of content density. Much effort in developing technical methods and personnel resources is needed to aggregate the geospatial reference data in a way that allows for a solid and quality proofed geo-referencing of geospatial thematic data (e.g. statistical analysis or environmental indicators). For improving this unsatisfactory situation, BGeoRG instructs BKG to evolve its competence and

capability to a central geospatial reference data and information service provider for the federal administration. Furthermore, the BGeoRG assigns BKG to advise the Federal Government in all questions related to geodesy and geospatial information and it safeguards the relevant German interests at the European and international level, e.g. within the United Nations.

Focusing on the geographical names standardization and maintenance process, the following institutions are involved in Germany:

- communes, counties, Länder (for populated places and administrative units)
- Länder survey administrations (for geographic regions/landscapes)
- federal and state hydrographic administration (for hydrographic features)
- federal and Länder transport administrations (for streets and railways)
- StAGN – Permanent Committee on Geographical Names (responsible for orthographic rules and principles for place names)

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

The following information related to finances, administration, legal issues and human resources only refers to the federal level. All tasks for the maintenance of the national geographical names database and its publication through web services and applications are conducted by the BKG, supported by the Permanent Committee on Geographical Names (StAGN).

In 1981 the first edition of the Gazetteer Federal Republic of Germany („West Germany“) was published

by the Institute for Applied Geodesy (IfAG; its name was changed to BKG in 1996). It comprised some 21.500 geographical names contained on the 1:500 000 General Topographic Map series. The data was available both in printed and digital form. At that time, it was the first gazetteer in Germany which had been prepared in accordance with the United Nations resolutions.

After the German reunification in 1990 this dataset was supplemented by the geographical names of the five new Länder (federal states) Mecklenburg-West Pomerania, Brandenburg, Saxony-Anhalt, Thuringia, Saxony and the eastern part of Berlin. The database was extended with the records of some 25,000 additional names rendered on the 1:250 000 general topographic map series. Since then, the national geographical names database called "Geographical Names of Germany" (GN-DE) has been developed and maintained by BKG. Today, it 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).

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	grosesse ewz
DEBKGGND000001IUD	276		142539	511051	220	67 42131

Endonym					
nnid	name	geschlecht	sprache_id	status_id	id
DEBKGGND000001IUD	Bautzen	1	1		
DEBKGGND000001IUD	Budyšin	2	1		

DLMLink			
nnid	ui_id	modell_id	layer_id
DEBKGGND000001IUD	DEBKGD200000WJK	1	12
DEBKGGND000001IUD	DEBKGD2000015N8	1	13

PLZ	
nnid	plz
DEBKGGND000001IUD	02625

Figure 3-1 GN-DE database table showing that one GNOBJekt may be associated to one or more geographical names (endonyms) in different languages. The example shows the GNOBJekt associated with 'Bautzen' (in German language) and 'Budyšin' (in Sorbian language). Both names are treated equally in the GN-DE and have equal official status.

At present the GN-DE database contains about 165,000 entries including more than 700 names in the languages of the national Sorbian and Frisian linguistic minorities. About 27,000 entries indicating the gender of hydrographic features are contained as well. The administration and maintenance tasks of maintaining the GN-DE are conducted at BKG in Frankfurt am Main, whereas the web service and applications are facilitated through the 'Central Service and Distribution Centre for Geoinformation' at the BKG department in Leipzig [1].

Search geographical names

Here you can search the geographic names contained in the dataset GN-DE (Geographische Namen Deutschlands / Geographic Names of Germany) online. For more information on our WFS GN-DE see our documentation.

Find a Name

Berlin

Name (min.2 characters)

* for arbitrary characters)

Map Presentation and Search



© GeoBasis-DE / BKG 2016
Search in small map boxes only. At most 100 objects are shown.

Name	Object Type	LAT	LON	Height	Height_calc	Area	Population	Language	Status	Map
Berlin	AX_Ortslage	13°24'20"	52°31'4"	35 m	34 m	892 km ²	3421829	deutsch	amtlich	
Berlin		10°26'52"	54°2'10"		37 m		474	deutsch	amtlich	
Berlin	AX_Gemeinde	13°26'47"	52°31'4"				3421829	deutsch	amtlich	
Berlin	AX_Bundesland	13°26'47"	52°31'4"				3421829	deutsch	amtlich	
Berlin	AX_KreisRegion	13°26'47"	52°31'4"				3421829	deutsch	amtlich	
Berlin	AX_Verwaltungsgemeinschaft_ATKIS	13°23'21"	52°30'39"				3421829	deutsch	amtlich	

Figure 3-1 Screenshots from the German geographical names application based on a Gazetteer Web Feature Service (WFS)

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 [3].

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

The following financial and human resources can be roughly estimated for maintaining the national geographical names programme by BKG. The figures can vary significantly according to the time and resources needed for the adaptation of the database, the services and applications to new technical developments or to the services oriented architecture in the organization.

BKG personnel costs: about 90,000 Euros / year;

- Office of Permanent Committee on Geographical Names, 1 person, full time, about 40,000 Euros / year
- Coordination of names activity at BKG, 2 persons, 30 %, about 10,000 Euros / year
- Technical staff for administration of databases, 1 person, 50 %, about 20,000 Euros / year
- Technical staff for administration of web (gazetteer) service, 1 person, 50 %, about 20,000 Euros / year

BKG investments: about 20,000 Euros / year

- Database (hard- and software), 2 databases, updates per year for 5,000 Euros / year
- Server (hard- and software), 1 server update per year for 5,000 Euros / year
- IT-specific training, 2 persons, about 10,000 Euros / year
- Soft- and hardware amortization is different for server or database and varies between 3-5 years.

3.4 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 'avoidable' 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 Centre 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) [4]. 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

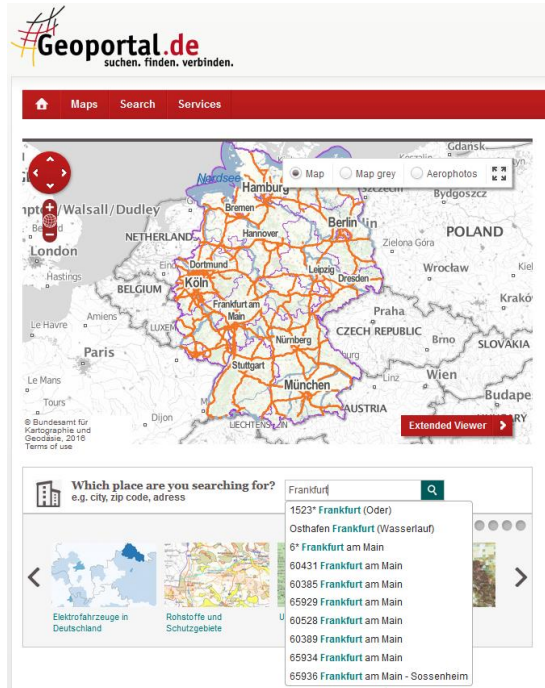


Figure 3-2 The German national geographical names database (GN-DE) published as a web service which is part of the national spatial data infrastructure (GDI-DE) and visualised through the Geoportal application (Geoportal.de)

for places and addresses of the Geoportal.de is envisaged, but not realized as yet (2016).

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' 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 3-3 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 [5]. For

geometry handling and implementation of OGC-standards the GeoTools library is used [6]. 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.

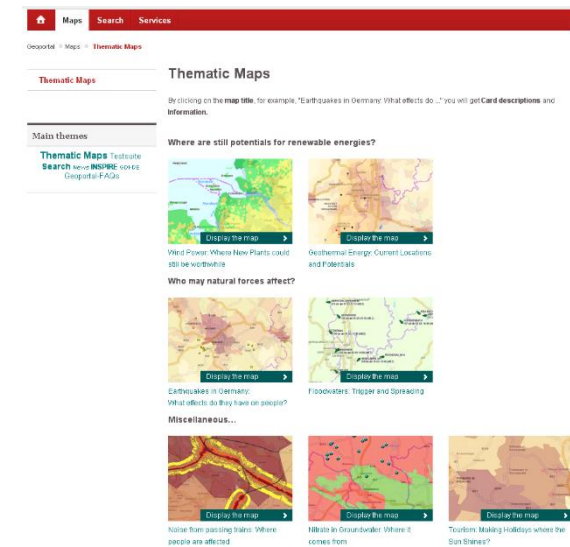


Figure 3-4 Geoportal.DE as a showcase for Spatial Data Infrastructure (SDI) through featured maps

In due course, a map showing the physio-geographic regions of Germany [5] combined with touristic information will be integrated. 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).

3.5 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 [8]. Thus, BKG has been one of the first institutions in Europe providing an INSPIRE-compliant national geographical names data service. According to the INSPIRE schedule EU countries are only obliged to realize this by the year 2017.

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.

INSPIRE GN data schema:

	GN data of BKG
mandatory	- name(s) (different spellings)
	- geometry
	- unique object identifier
'voidable'	- feature type (feature type classification)
	- language (3-letter code ISO 639-3 or -5)
	- nameStatus (official, standardised, historical, other)
	- link to related SpatialObject
	- script (4-letter code ISO 15924)
	- nativeness' (endonym, exonym)
	- transliterationScheme
	- grammatical gender (masc., fem., neuter, common)
	- grammatical number (singular, plural, dual)
	- pronunciation (IPA, audio files)
	- sourceOfName
- typeLocal	
- lifeCycleInfo (start/end of the feature/object in the original DB)	
	* ...

Figure 3-5 INSPIRE Data schema fulfilled by BKG

3.6 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 [9].

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 [10].

3.7 Further discussion on logistics, production time needed and necessary upkeep costs

The surveying and cadastral administrations of the Länder, the statistical offices and the Federal Statistical Office, private and academic institutions for language and culture etc. are all potentially data-holding authorities for INSPIRE (in accordance with the German law 'Geodatenzugangsgesetz – GeoZG') for geographical names data, though so far with highly different names density.

The financial and human resources roughly estimated by BKG as necessary for maintaining the national geographical names programme amounted to about € 110,000 per year for both personnel and investments. The figures vary according to the time and resources needed for the adaptation of the database, the services and applications to new technical developments or to the services oriented architecture in the organization. Considering that most of the German Länder might

invest a similar small amount in logistics, the maintenance of geographical names database and the development and up keeping of services and applications, about the overall yearly costs would amount to € 1,8 million.

Considering that the data content for GN-DE is based on a map scale 1: 250,000 a cooperation of the Federal Government with the Länder– (the latter are responsible for data based on maps at scales 1: 25,000 – 1: 10,000) is envisaged in order to provide a unique source for geographical names data in Germany and increase the content of the national web service and integrate names data taken from large scale maps.

The Länder do not and will not add any toponymic attributes like status or language to the geographical names in their databases, as this has never been part of their tasks. The geographical names data content is based on the ATKIS programme and consists of the geographical names as an associated attribute to the object/feature and by that inherits only its feature type and its geometry. Thus, joining the names datasets of BKG and of the Länder and so providing one unique geographical names data source for Germany would be beneficial to everyone – including data users. Furthermore, the INSPIRE data provision could be facilitated by BKG for the federal and the Länder administrations with one national web service.

3.8 References

- [1] Central Service and Distribution Centre for Geoinformation at BKG, www.geodatenzentrum.de, last accessed 09/2016
- [2] Working Committee of the Surveying Authorities of the Länder of the Federal Republic of Germany <http://www.adv-online.de>, last accessed 09/2016
- [3] Geographische Namen Deutschland – GN-DE, the (gazetteer) WFS service (wfs_gnde), http://www.geodatenzentrum.de/xml/WFS_gnde.xml, last accessed 09/2016
- [4] Geographische Namen 1: 250,000 (GN250), documentation of the dataset (in German only), <http://www.geodatenzentrum.de/docpdf/gn250.pdf>, last accessed 09/2016
- [5] Apache Lucene: <http://lucene.apache.org>, last accessed 09/2016
- [6] <http://geotools.org/>, last accessed 09/2016
- [7] 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)
- [8] 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), [\[lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:323:0011:0102:DE:PDF\]\(http://lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:323:0011:0102:DE:PDF\), last accessed 09/2016](http://eur-</div><div data-bbox=)

[9] 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 09/2016

[10] Open Data - Freie Daten und Dienste des BKG: http://www.geodatenzentrum.de/geodaten/gdz_rahme_n.gdz_div?gdz_spr=eng&gdz_akt_zeile=5&gdz_anz_zeile=0&gdz_user_id=0, last accessed 09/2016