

Session:

Database creation (theory)

v1.0

Pier-Giorgio Zaccheddu

Federal Agency for Cartography and Geodesy (BKG)

Richard-Strauss-Allee 11

60598 Frankfurt am Main

Tel.: +49 69 6333 305 Fax: +49 69 6333 441

Email: pier.zaccheddu@bkg.bund.de

Content

Preparing the names database – General issues

- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

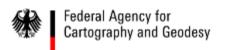
- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Data model vs. database modelling

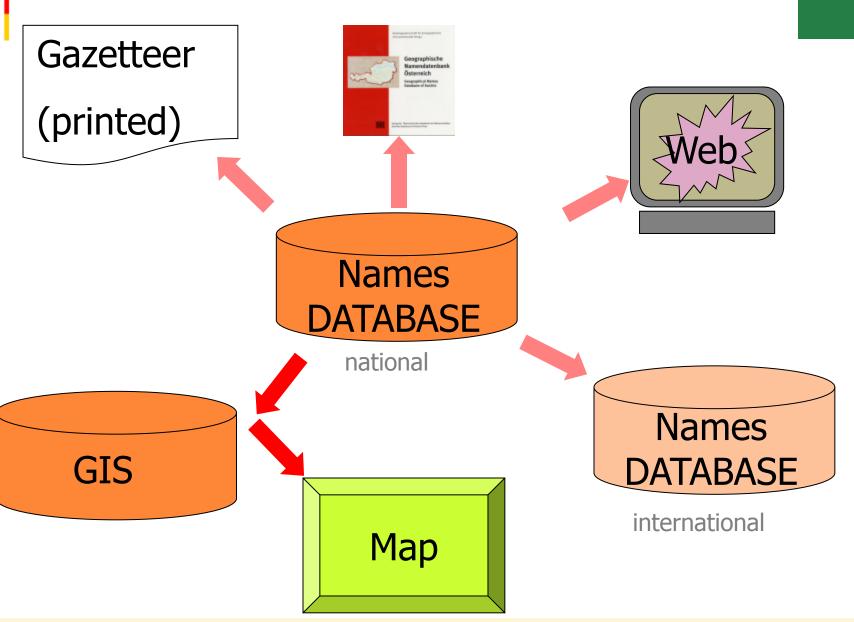
Objectives for the 'hands on' experience for creating a database I don't want to...

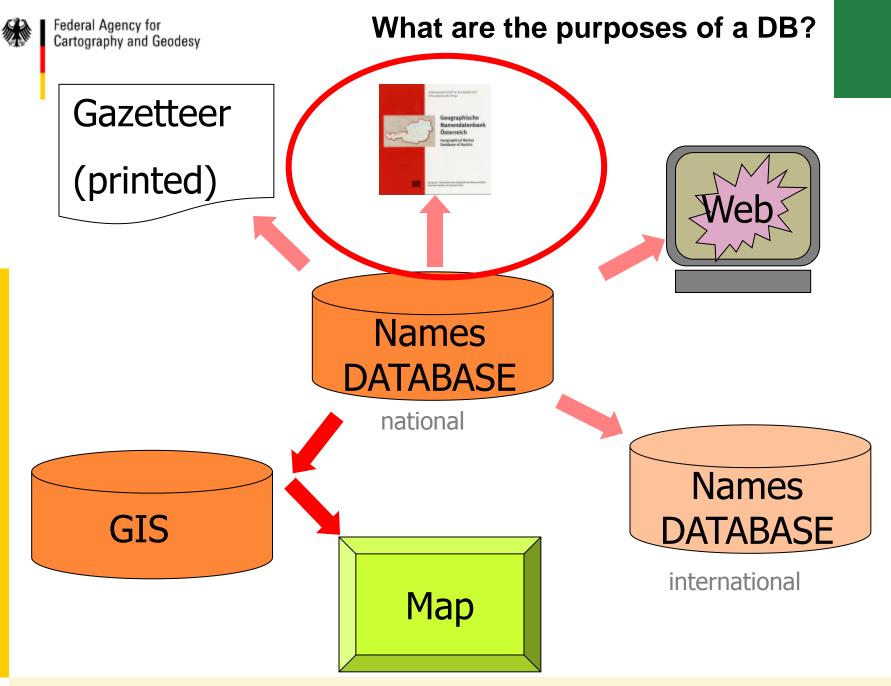
- invest too much time in gaining technical IT knowledge
- understand the technical database management in detail
- use Geographic Information System (GIS) in future work
- technically understand Web services and Web applications

I do want to...

- understand a database model in principle
- discuss with and explain to my technical IT staff the use cases and database content
- explain the database model and its concept to my staff









What are the purposes of a DB? Example: Netherlands

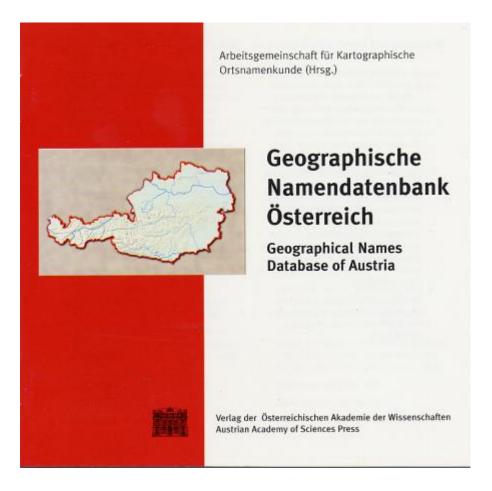


Geographical
Names Database:
Netherlands

Offline DB



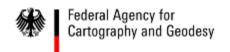
What are the purposes of a DB? Example: Austria



Geographical Names Data Base: Austria

Offline CD-ROM

Top Train Course, Rio de Janeiro, Brazil

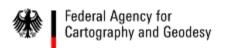


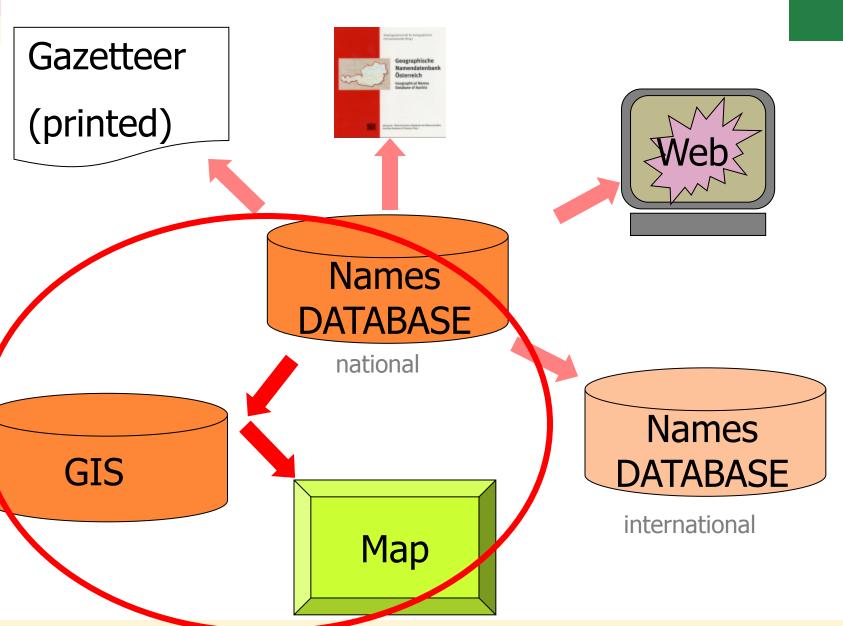
What are the purposes of a DB? Example: Germany



Example of a topographic data file with integrated names data base: Germany

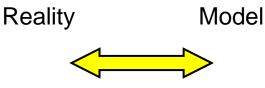
Offline CD-ROM













- Where am I?
- Where do I find...?
- Where is the next...?
- How do I get to...?
- How far is it to...?
- Where does this way lead to?

Data models

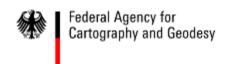
↓ consist of

feature definitions + relations

↓ including

spatial reference to points,
locations, areas or regions as specific feature

- → necessary information is called geo(graphic) information
- → 80% of all information is estimated to be spatially referenced

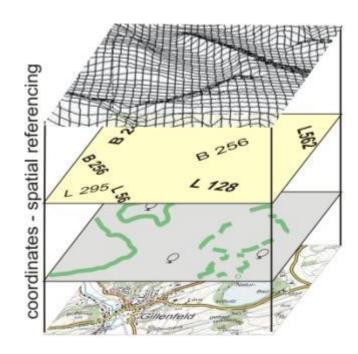


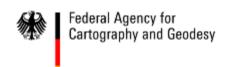
Geographic Information System (GIS)

a GIS is an organized collection of computer hardware, software, (geographic) data and personnel.

→ designed to capture, store, update, manipulate, analyse and display all forms of geo referenced information.

(see www.GIS.com)





Geo(graphic) information:

Information that is referenced to the earth's surface, whether by coordinates (direct referencing) or by identifiers such as addresses or postal codes or geographical names (indirect referencing).

Geo(graphic) data / spatial data:

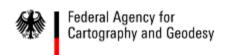
Computer-readable geo information

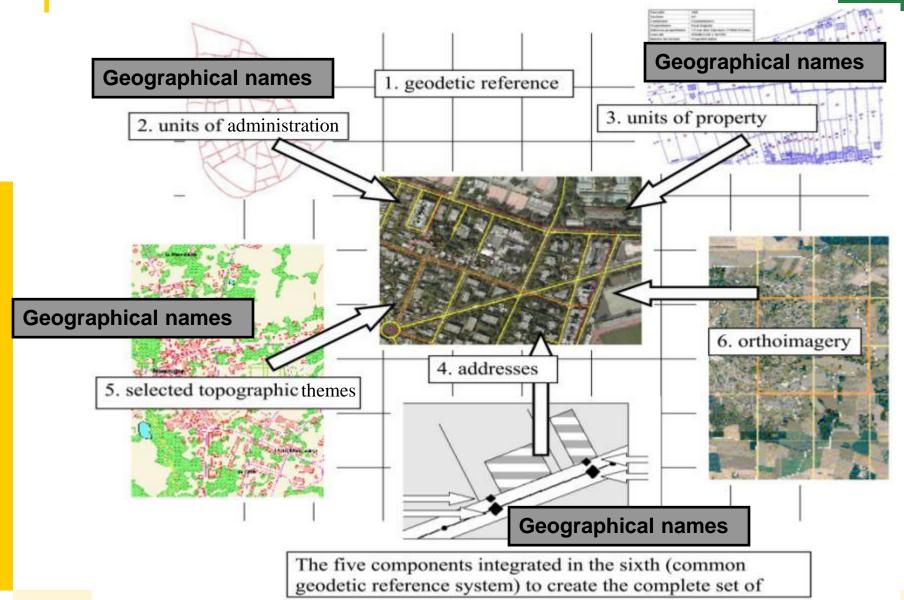
Vector data model (feature data)

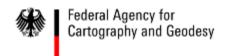
points, lines and polygons (areas)

Raster data model (coverage data)

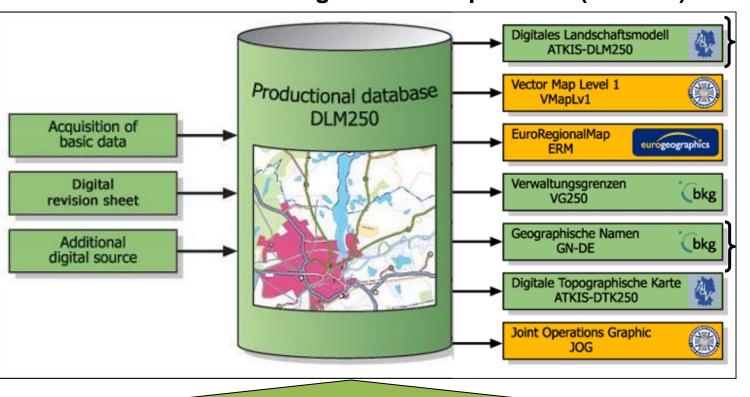
gridded data (scanned maps, satellite images, orthophotos)







BKG's produkt line - derived from the Digital Landscape Model (DLM250)

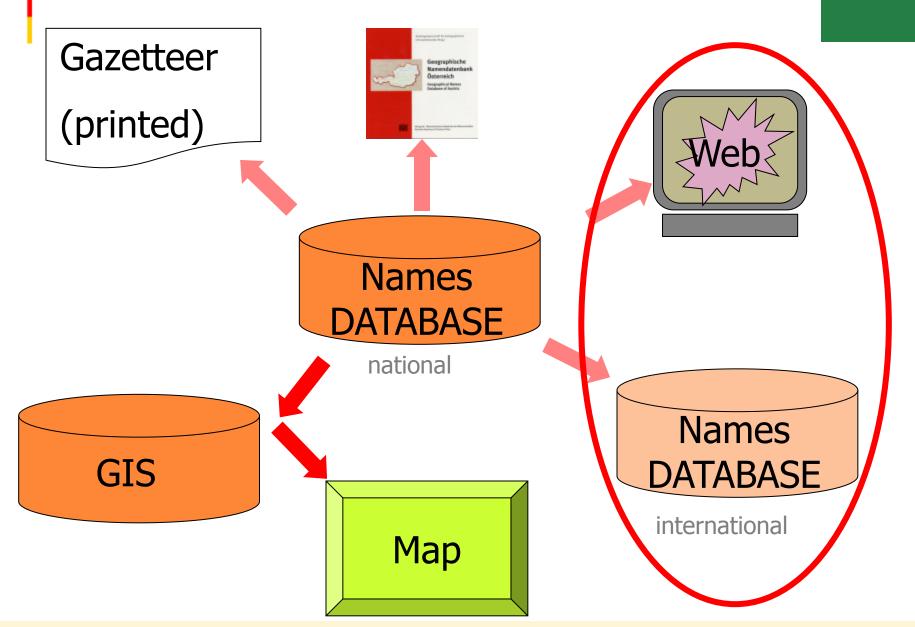


DLM for the AdV product line

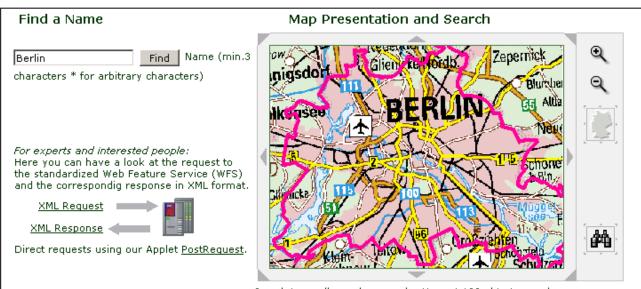
GN dataset for the national gazetteer service

Geographical names data set **GN-DE** is essential part of the productional database and is unambiguously linked to the **spatial objects/features** through **unique identifiers (UIDs)**





The national gazetteer service – search for geographical names of Germany...



Gazetteer service based on the dataset **GN-DE**

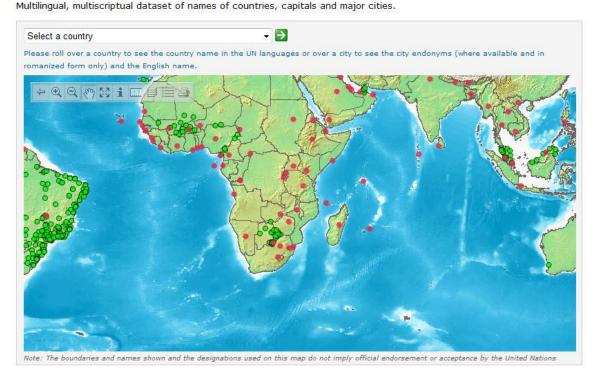
Search in small map	boxes	only.	At most 100	objects	are shown.
---------------------	-------	-------	-------------	---------	------------

Name	Object Type	Postcode	LAT	LON	Height	Area	Population	Language	Status	Мар	
Berlin	Ortslage	23823	10°26'56"	54°02'33"				deutsch	amtlich		
Berlin	Ortslage	10115 🔽	13°24'37"	52°31'19"	34 m	892 km²	3387828	deutsch	amtlich		
Berlin	Siedlung		13°24'37"	52°31'19"				deutsch	amtlich	Fa	
Berlin	Verwaltungseinheit					- 3	as gra	aphic	al W	eb/	Application
Berlin	Verwaltungseinheit					_ 4	ac OG	CW	aptlick	00	turo Sorvico (MES)
Berlin	Verwaltungseinheit						as OC		an L	ea	ture Service (WFS)

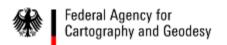
The international UNGEGN database – search for geographical names of countries and capitals



Geographical names | FAQ | Feedback | Contact

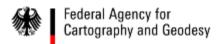


http://unstats.un.org/unsd
/geoinfo/geonames/



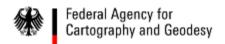


Multilingual, multiscriptual dataset of names of countries, capitals and major cities. Geographical Names Database | About | FAQ | Feedback | Contact Beta version Brazil Roll over a country to display the country name in the six UN languages. Point to a city to see the city endonyms (where available and in romanized form only) and the English name. Capital cities are shown in red, other cities are shown in green. Note: Depending on the computer settings, some names might not be able to be displayed properly.

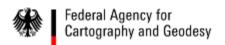


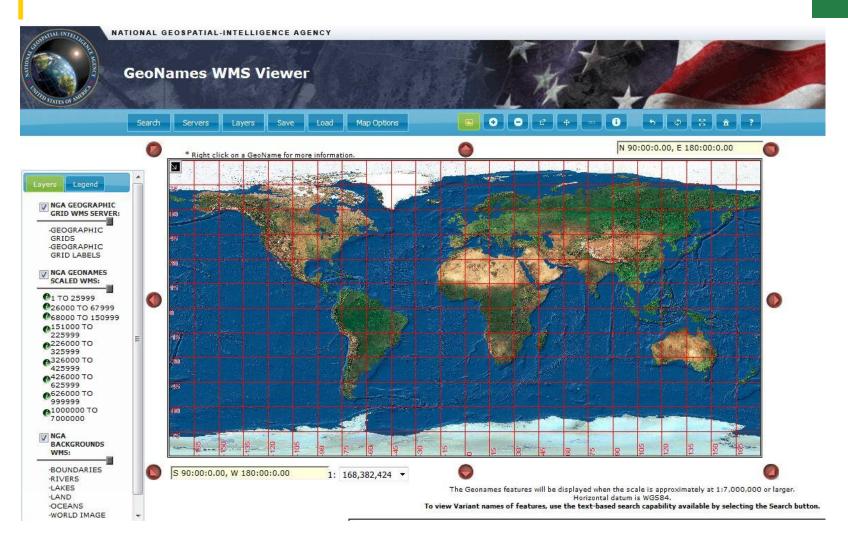






Lat: -32.03 Long: -52.10	Rio Grande	
Endonym		
Language	City name	Source
Portuguese	Rio Grande 🕬	UNGEGN
Lat: -31.77 Long: -52.34	Pelotas	
Lat: -31.33 Long: -54.11	Bagé	
Lat: -30.08 Long: -51.02	Viamão	
Lat: -30.03 Long: -51.23	Porto Alegre	
Lat: -29.99 Long: -51.08	Alvorada	
Lat: -29.95 Long: -51 <mark>.</mark> 09	Cachoeirinha	
Lat: -29.94 Long: -50.99	Gravataí	
Lat: -29.92 Long: -51.18	Canoas	
Lat: -29.84 Long: -51.14	Sapucaia do Sul	
Lat: -29.76 Long: -51.15	São Leopoldo	
Lat: -29.75 Long: -57.09	Uruguaiana	
Lat: -29.72 Long: -52.43	Santa Cruz do Sul	
Lat: -29.68 Long: -53.81	Santa Maria	





Search for geographical names through the GEOnet Names Server (GNS) http://geonames.nga.mil/gns/html/, last accessed 04/2017



Gazetteer for Scotland	http://www.scottish-places.info/
Gazetteer of Japan	http://www.gsi.go.jp/common/000042053.pdf
Magyarország Helységnévtára, 2012 - The Detailed Gazetteer of Hungary, 2012	http://www.ksh.hu/apps/cp.hnt2.gazetteer
ScotlandsPlaces, Historical place name gazetteer	http://www.scotlandsplaces.gov.uk/
The National Gazetteer of Wales	http://homepage.ntlworld.com/geogdata/ngw/places.htm
Composite Gazetteer of Antarctica	https://www1.data.antarctica.gov.au/aadc/gaz/scar/search.cfm
Gazetteer of British Place Names	http://www.gazetteer.org.uk/index.php
Getty Thesaurus of Geographic Names Online	http://www.getty.edu/research/tools/vocabularies/tgn/
Bavarikon, Ortssuche	http://www.bavarikon.de/places
Gazetteer of Australia - Geoscience Australia, Place Name Search	http://www.ga.gov.au/place-names/index.xhtml
Geographic Names Information System	http://geonames.usgs.gov/apex/f?p=136:1:8584754151386
Canadian Geographical Names Data Base - CGNDB	http://www.nrcan.gc.ca/earth-sciences/geography/place-names/search/9170
NWT Place Names Database	http://www.pwnhc.ca/cultural-places/geographic-names/database-of-nwt-geographic-r
North Dakota Place Names	http://www.webfamilytree.com/North Dakota Place Names/index.htm
Directory of Towns, Villages and Hamlets Past and Present of Missouri	http://thelibrary.org/lochist/moser/search.html
A Gazetteer of Cities, Towns, Villages, Mountains, Hills, Rivers, and other Place names with their location, la	https://www.placenames.com/
GeoTWAIN 2.0	http://geotwain.de/
Openstreetmap Nominatim	https://nominatim.openstreetmap.org/
Crimean Tatar Place Names	http://www.iccrimea.org/place/placenames.html
Geographic Names Informations System, Query Form for Antarctica	http://geonames.usgs.gov/apex/f?p=gnispq:1:0::NO:1:P1 SHOW ANTAR%2CP1 SHOW A
Interaktyvus Rytų Prūsijos žemėlapis II IRPŽ II - Interaktive Karte von Ostpreußen	http://prusija.lki.lt/
Antarctic Place Names	https://www1.data.antarctica.gov.au/aadc/gaz/
Registar Geografskih Imena - Gazetteer of Geographical Names	http://cgn.dgu.hr/name-search/
Registar Geografskih Imena Nacionalnih Manjina Republike Hrvatske - Gazetteer of national minorities of th	http://www.dgu.hr/assets/uploads/Dokumenti/Poslovne%20info/Registar%20geog%20i
Autoriserede stednavne i Danmark - List of authorized Danish place-names	http://www.stednavne.info/VisKort/Popupmap.aspx
Maanmittauslaitos - National Land Survey of Finland	http://kansalaisen.karttapaikka.fi/kartanhaku/paikannimihaku.html?e=406643&n=71951
Géoportail, le portail des territoires et des citoyens - Geoportal Frankreich	http://www.geoportail.gouv.fr/accueil
Geographische Namen Deutschlands GN-DE	http://www.geodatenzentrum.de/geodaten/gdz rahmen.gdz div?gdz spr=deu&gdz ak
Historische Ortsnamen	http://www.geodatenzentrum.de/geodaten/gdz rahmen.gdz div?gdz spr=deu&gdz ak

→ Session Names Servers...on Friday, May 19

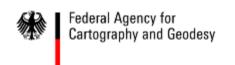
Content

Preparing the names database – General issues

- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Data model vs. database modelling



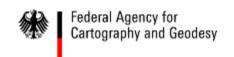
Use cases for Web applications using names data services



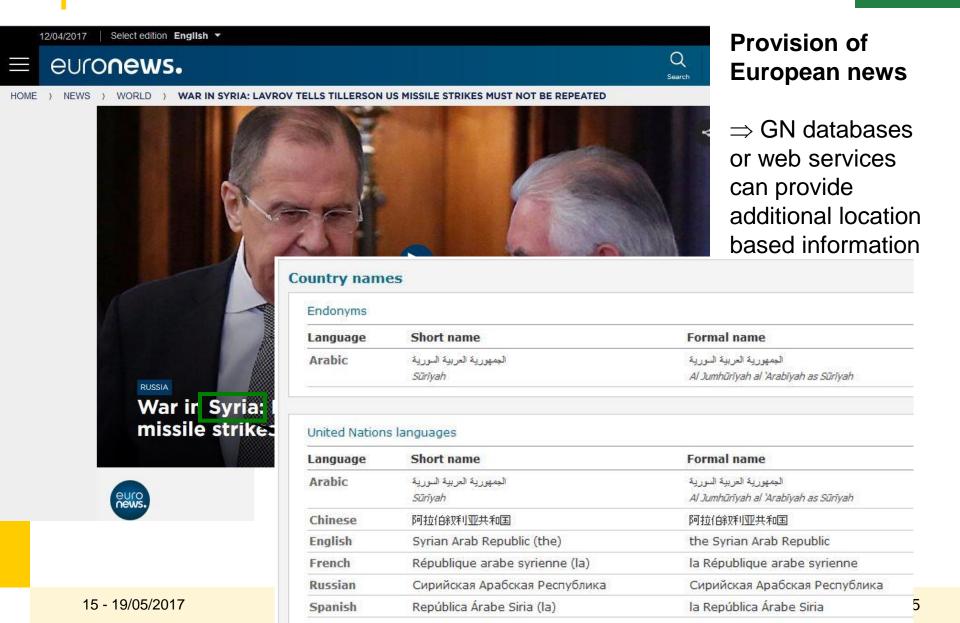
International purchase of properties with Internet property services

http://www.viviun.com/

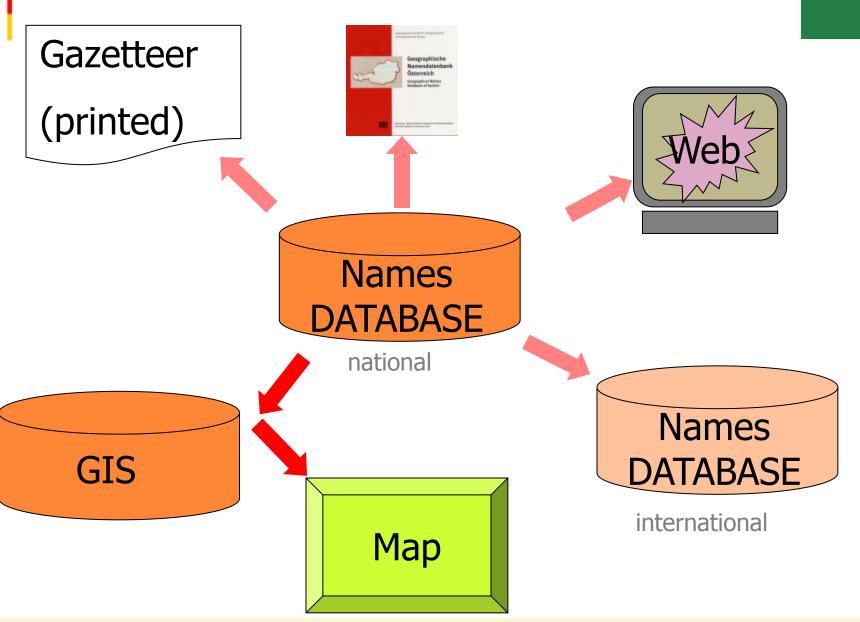
⇒ Geographical names databases or web services can provide additional location based information



Use cases for Web applications using names data services









Eighth United Nations Conference on the Standardization of Geographical Names, 2002

Resolution VIII / 6

Integration of Geographical Names Data into National and Regional Spatial Data Infrastructures

The Conference,

. . . .

Recommends,

that standardized geographical names data should be considered in the establishment of national and regional spatial data infrastructures and included in their constructions.

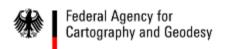


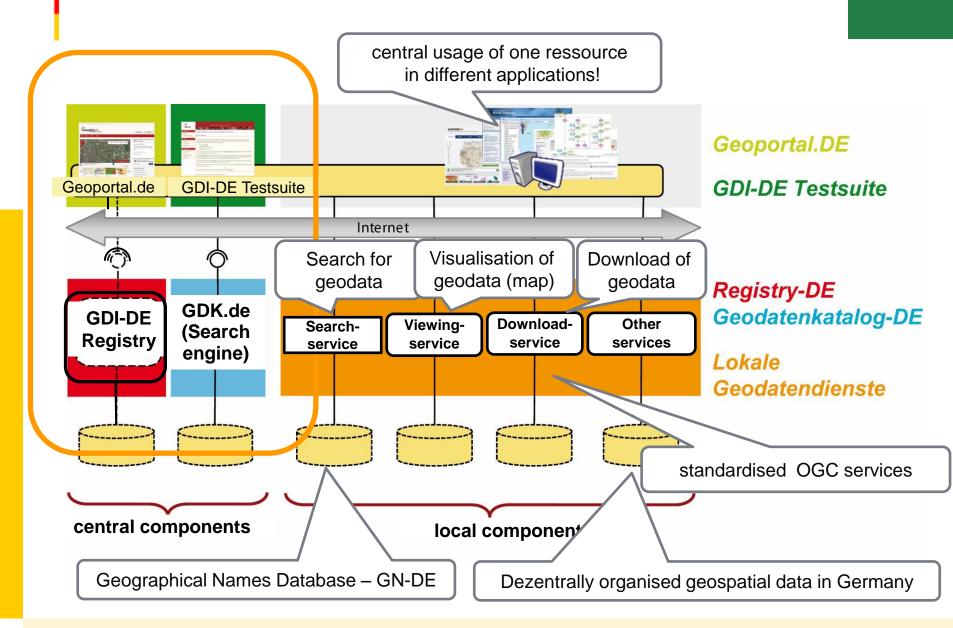
A spatial data infrastructure (SDI) is a data infrastructure implementing a framework of geographic data, metadata, users and tools that are interactively connected in order to use spatial data in an efficient and flexible way.

Another definition is:

[...] the technology, policies, standards, human resources, and related activities necessary to acquire, process, distribute, use, maintain, and preserve spatial data.

The White House - Office of Management and Budget (2002) Circular No. A-16 Revised, August 19, 2002









The German national geographical names database (GN-DE) published as a web service as part of the national spatial data infrastrucuture (GDI-DE) and visualized through the Geoportal application (Geoportal.de)

Content

Preparing the names database – General issues

- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Data model vs. database modelling





What do you have to consider (or to know) before you start with creating a database?

Different computer support different characteristics...

Data Capture

Data Storage



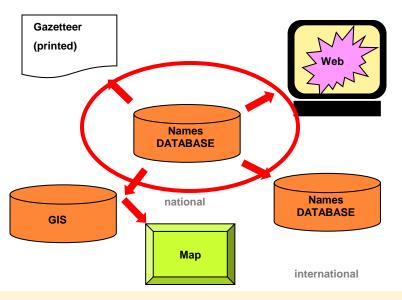
Data Processing



Data Output

from cards, lists, books, fieldwork, maps, other databases,

in a database



1. Operating system

An operating system (OS) is a <u>software program</u> that manages the <u>hardware</u> and <u>software</u> resources of a <u>computer</u>. The OS performs basic tasks, such as controlling and allocating memory, prioritizing the processing of instructions, controlling input and output devices, facilitating networking, and managing files.

Examples:

- Microsoft Windows (proprietary)
- Linux (open source = free of charge and usage)



2. Database

One possible definition is that a database is a collection of <u>records</u> stored in a computer in a systematic way, so that a <u>computer program</u> can consult it to answer questions.

Examples:

- Microsoft Access, Oracle, etc. (proprietary)
- MySQL, PostGreSQL/PostGIS, SQLite, etc. (open source)

3. Geographic Information System (GIS)

A **geographical information system** (GIS) is a system for creating, storing, analyzing and managing spatial data and associated attributes.

In a more generic sense, GIS is a tool that allows users to create interactive queries (user created searches), analyze the spatial information, and edit data.

Examples:

- ESRI ArcGIS, ViewMap, etc. (proprietary)
- Quantum GIS, etc. (open source)



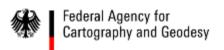


e.g. Codepages!!!

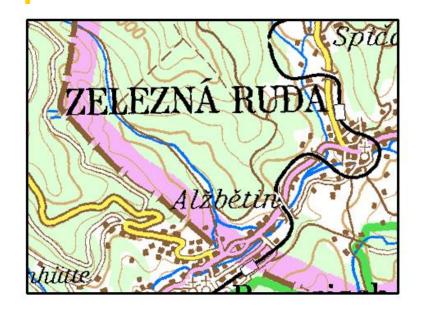
ISO8859 character set or Unicode? in (Microsoft Windows) operating system

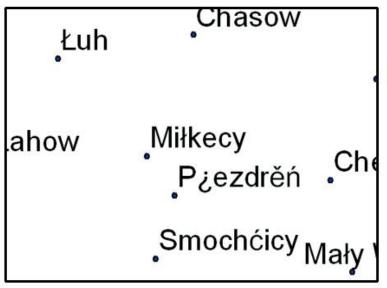
Different interpretation of characters possible in

- Database products: MS Access or similar
- GIS software: ESRI ArcGIS or similar



Characteristics of databases







Nowe £a
Aleksandrów £ódzki
Konstantynów £ódzki

Content

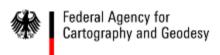
Preparing the names database – General issues

- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Data model vs. database modelling





Operating System:

Microsoft Windows?

Linux? (open source)

Database:

Microsoft Access?

MySQL? PostGreSQL/PostGIS? (open source)

UNECA GeoNyms application? (open source)

Geographical Information System

ESRI ArcGIS?

Quantum GIS? (open source)

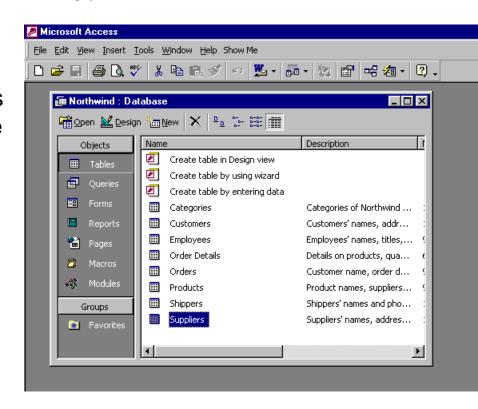
Realisation of databases



Database:

Microsoft Access? (proprietary)

- Microsoft Access is a pseudorelational database management system from Microsoft that combines the relational Microsoft Jet Database Engine with a graphical user interface and software-development tools.
- Access stores data in its own format based on the Access Jet Database Engine. It can also import or link directly to data stored in other applications and databases.
- Website: http://office.microsoft.com/en-us/access/

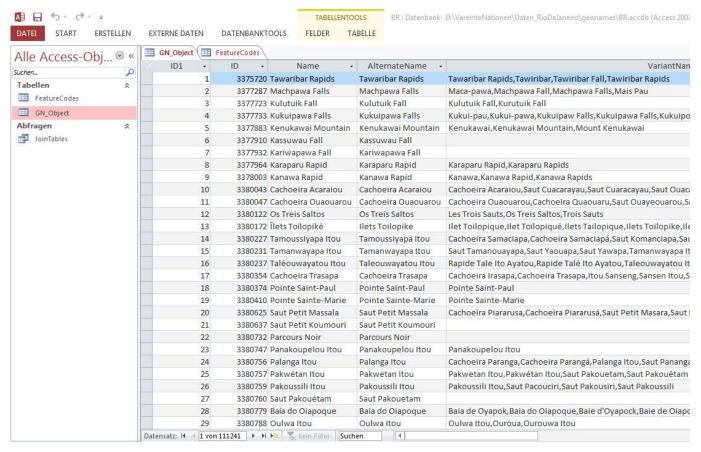




Realisation of databases

Database:

Microsoft Access? (proprietary)







Database:

MySQL? PostGreSQL/PostGIS? (open source)

My SQL "The world's most popular open source database"

- MySQL is a relational database management system (RDBMS) that runs as a server providing multi-user access to a number of databases.
- The MySQL development project has made its source code available under the terms of the GNU General Public License, as well as under a variety of proprietary agreements.
- MySQL is a popular choice of database for use in web applications.
- Website: http://www.mysql.com/





Database:

MySQL? PostGreSQL/PostGIS? (open source)

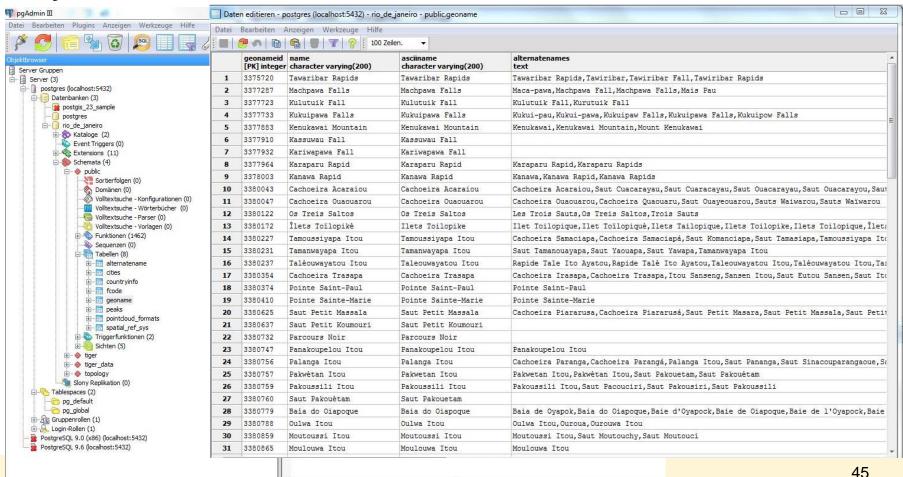
PostGreSQL / PostGIS (open source)

- PostgreSQL, often simply Postgres, is an object-relational database management system (ORDBMS).
- PostgreSQL is not controlled by any single company a global community of developers and companies develops the system.
- PostGIS adds support for geographic objects to the PostgreSQL object-relational database. In effect, PostGIS "spatially enables" the PostgreSQL server, allowing it to be used as a backend spatial database for geographic information systems (GIS), much like ESRI's SDE or Oracle's Spatial extension.
- Website: http://www.postgresql.org/
 http://postgis.refractions.net/

Realisation of databases

Database:

MySQL? PostGreSQL/PostGIS? (open source)





Geographical Information System

ESRI ArcGIS? (proprietary)

ESRI Arc GIS

Esri® ArcGIS® facilitates collaboration and lets you author data, maps, globes, and models on the desktop and serve them for use on a desktop, in a browser, or in the field, depending on the needs of your organization.

- ArcGIS support and educational services
 consist of technical maintenance programs,
 software releases and updates, technical
 support, online support services, publications
 training, and consulting services.
- Website: http://www.esri.com/software/arcgis/index.ht ml



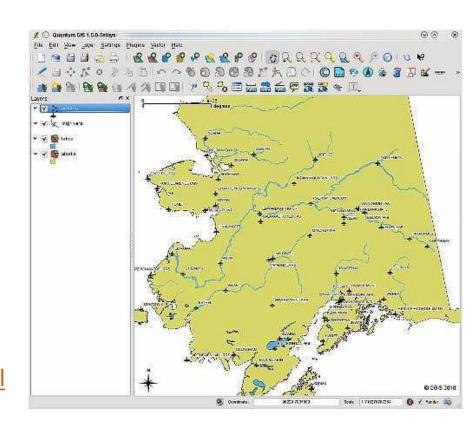


Geographical Information System

Quantum GIS? (open source)

Quantum GIS

- QGIS provides data viewing, editing, and analysis capabilities.
- Quantum GIS (QGIS) is a user friendly Open Source Geographic Information System (GIS) licensed under the GNU General Public License.
- QGIS runs on Linux, Unix, Mac OSX, and Windows and supports numerous vector, raster, and database formats and functionalities.
- Website: http://www.qgis.org/en.html



Open Source Software Collection

Open source software collection: OSGeo Live

URL: http://live.osgeo.org/en/index.html

about:

OSGeo-Live is a self-contained bootable DVD, USB thumb drive or Virtual Machine, that allows you to try a wide variety of open source geospatial software without installing anything. It is composed entirely of free software, allowing it to be freely distributed, duplicated and passed around.



Open your Mind through GIS: training in Open Source GIS



http://www.unige.ch/formcont/opengis.html

Content

Preparing the names database – General issues

- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Data model vs. database modelling

"UNGEGN schema" (proposal): elements

Field Name	Data Type	Description	Example
RecordID	Index	This number is assigned automatically by the system. Do not change it.	
UID	Number (long integer)	A <u>unique identifier</u> assigned to the name (or named feature). In our exercise, each group may start to count from 1 and add the group number multiplied by 100, e.g. 201 for the first name by group 2.	20001 might indicate e.g. 2000 for feature type 'Populated places'
Name	Text (50 char.)	Geographical name in Romanian	Oran
Latitude	Number (double)	Geographical Coordinates, in degrees decimal.	35,
Longitude	Number (double)		-0,5
FeatureCode	Text (50 char.)	Feature codes, from an existing feature code table	DDS
AdminUnit	Text (50 char.)	Name of the administrative unit where the name is situated in	Departement d' Oran, Department of Oran
Language	Text (50 char.)	Language of the name	
Description	Text (255 char.)	Field remarks, meaning of the name, language of the name, historical names if any	e.g. capital of a political entity
VariantName	Text (50 char.)	Enter variant names, if any e.g. Hungarian name	Wilaya d' Oran, Wilaya d' Oran
MapSheet	Number (long integer)	Reference to a map sheet in a topographic map series, e.g. 1:250.000	80
Source	Text (255 char.)	Source of the information on the name: - Informant - Interviewer	Mr. XY, old person at xyz, interview by group 1
Status	Text (50 char.)	The status of the name. In our case, the names are not yet approved by the Board.	not approved
Pronunciation	OLE-Object	Audio-files of the pronunciation of the geographical name	e.g *.wav -file
Location information	OLE-Object	Digital pictures of the location	e.g *.jpg – image file



The European (INSPIRE) GN schema: elements

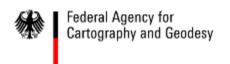
mandatory

- name(s) (text, spelling)
- geomety
- feature type
- unique identifier

- language {three letter codes from ISO 639-3 or -5}

- nameStatus {official, standardised, historical, other}
- link to relatedSpatialObject
- script (four letters codes defined in ISO 15924)
- nativeness {endonym, exonym}
- transliterationScheme
- grammatical gender {masc., fem., neuter, common}
- grammatical number {singular, plural, dual}
- pronunciation
- sourceOfName
- typeLocal
- lifeCycleInfo (begin/end of the object in the source DB)
- ...

'voidable'



INSPIRE priorities in Annex I-III and the selction of "Core Data" for Europe through UN-GGIM: Europe

Annex I

Coordinate Reference Systems

Geographical Grid Systems

Geographical Names

Administrative Units

Addresses

Cadastral Parcels

Transport Networks

Hydrography

Protected Sites

Annex II

Elevation

Land Cover

Ortholmagery

Geology

Annex III

Statistical units

Buildings

Soil

Land use

Human health and safety

Utility and governmental services

Environmental monitoring facilities

Production and industrial facilities

Agricultural and aquaculture facilities

Population distribution - demography

Area management/restriction/ regulation

Natural risk zones

Atmospheric conditions

Meteorological geographical features

Oceanographic geographical features

Sea regions

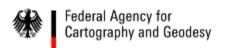
Bio-geographical regions

Habitats and biotopes

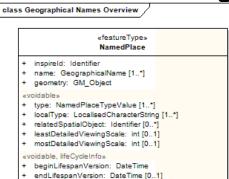
Species distribution

Energy resources

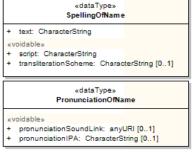
Mineral resources



INSPIRE Geographical Names – UML Schema

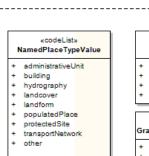






«codeList»

«codel ist»





A Named Place,

representing a real world entity referred to by a Geographical Name

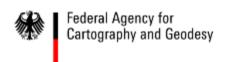
is associated with one or several Geographical Names,

i.e. proper noun applied to the feature

- "Athina" language = Greek nativeValue = Endonym
- "Athens" language = English (2)nativeValue = Exonym

and may have one or several Spellings of Name.

i.e. proper way of writing the name



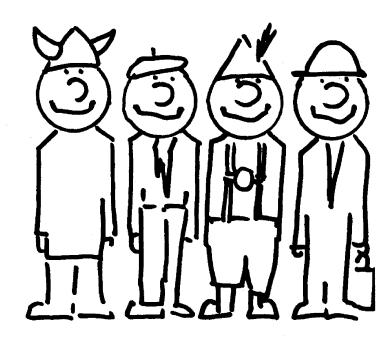
Standardization in Europe

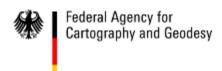
Not standardized and very different

Almost harmonized, but with national specialities









Thank you for your attention!