Federal Agency for Cartography and Geodesy

Session:

Preparing the names database

V1

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
- Example: UNECA initiative – GeoNyms desktop database application
- Data model vs. database modelling

What are the purposes of a DB?

Gazetteer

(printed)

Web

Names
DATABASE

national

GIS

Map

Names
DATABASE

international

What are the purposes of a DB?

Gazetteer

(printed)

Web

Names
DATABASE

national

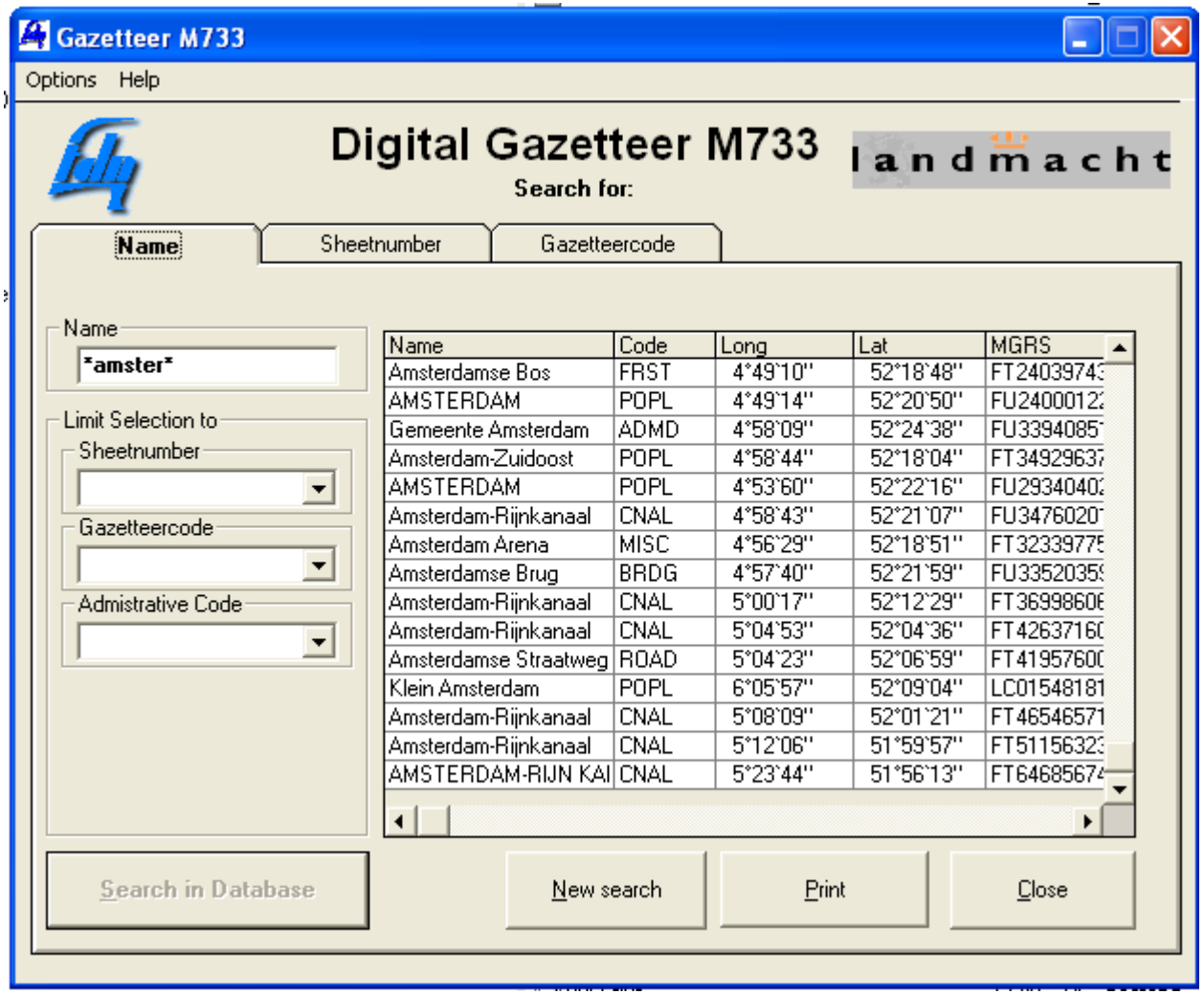
GIS

Map

Names
DATABASE

international

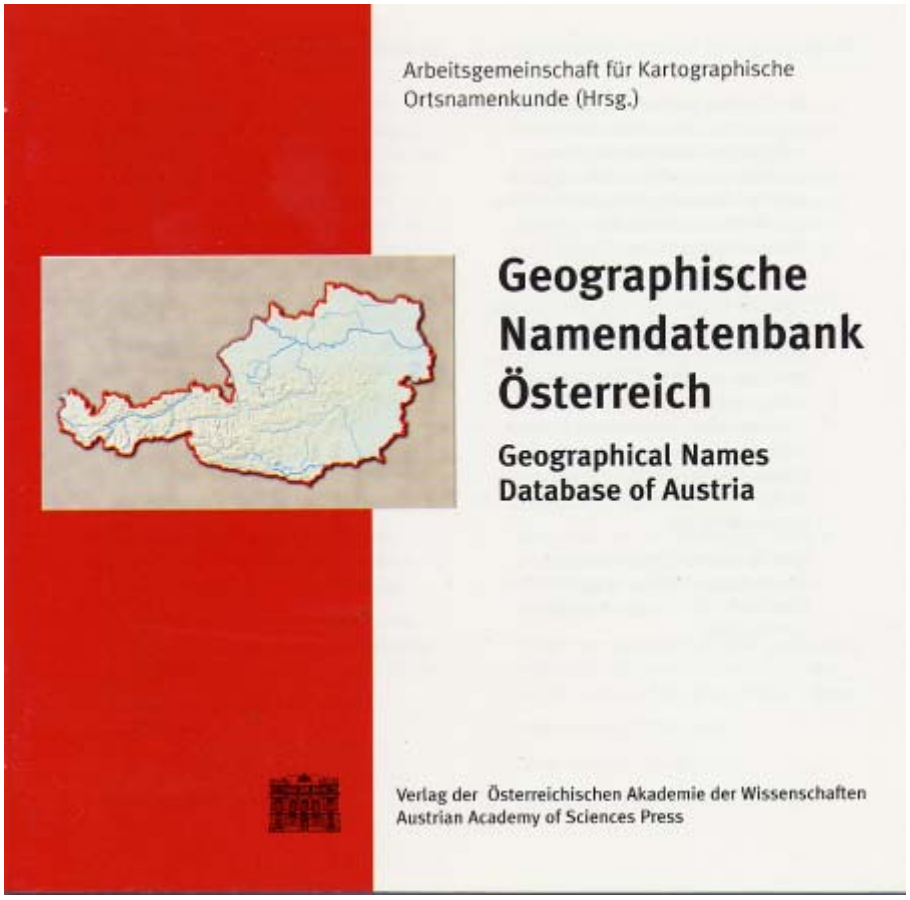
What are the purposes of a DB?
Example: Netherlands



Geographical Names Database: Netherlands

Offline DB

What are the purposes of a DB? Example: Austria



Geographical Names Data Base: Austria

Offline CD-ROM

What are the purposes of a DB? Example: Germany

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

Offline CD-ROM

What are the purposes of a DB?

Gazetteer

(printed)

Web

Names
DATABASE

national

GIS

Map

Names
DATABASE

international

Federal Agency for Cartography and Geodesy

Names database and Geographic Information System (GIS)

Reality ⟷ Model

- **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***

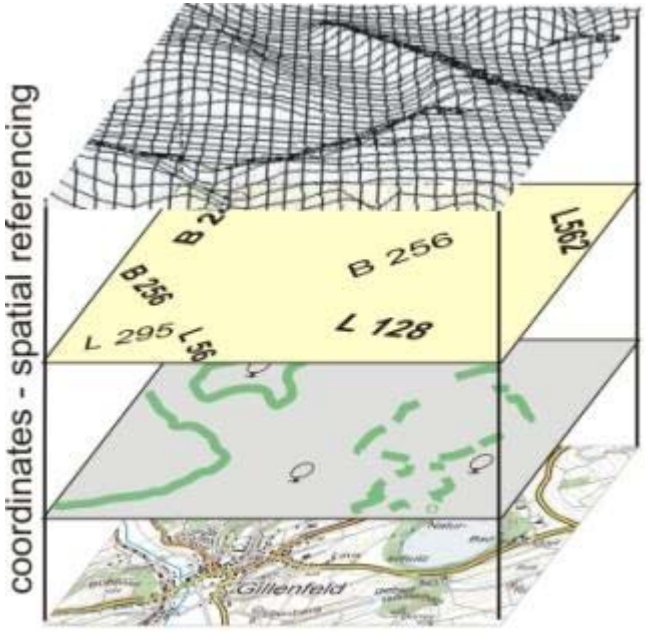
→ ***80% of all information is estimated to be spatially referenced***

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)



Names database and Geographic Information System (GIS)

Geo(graphic) information:

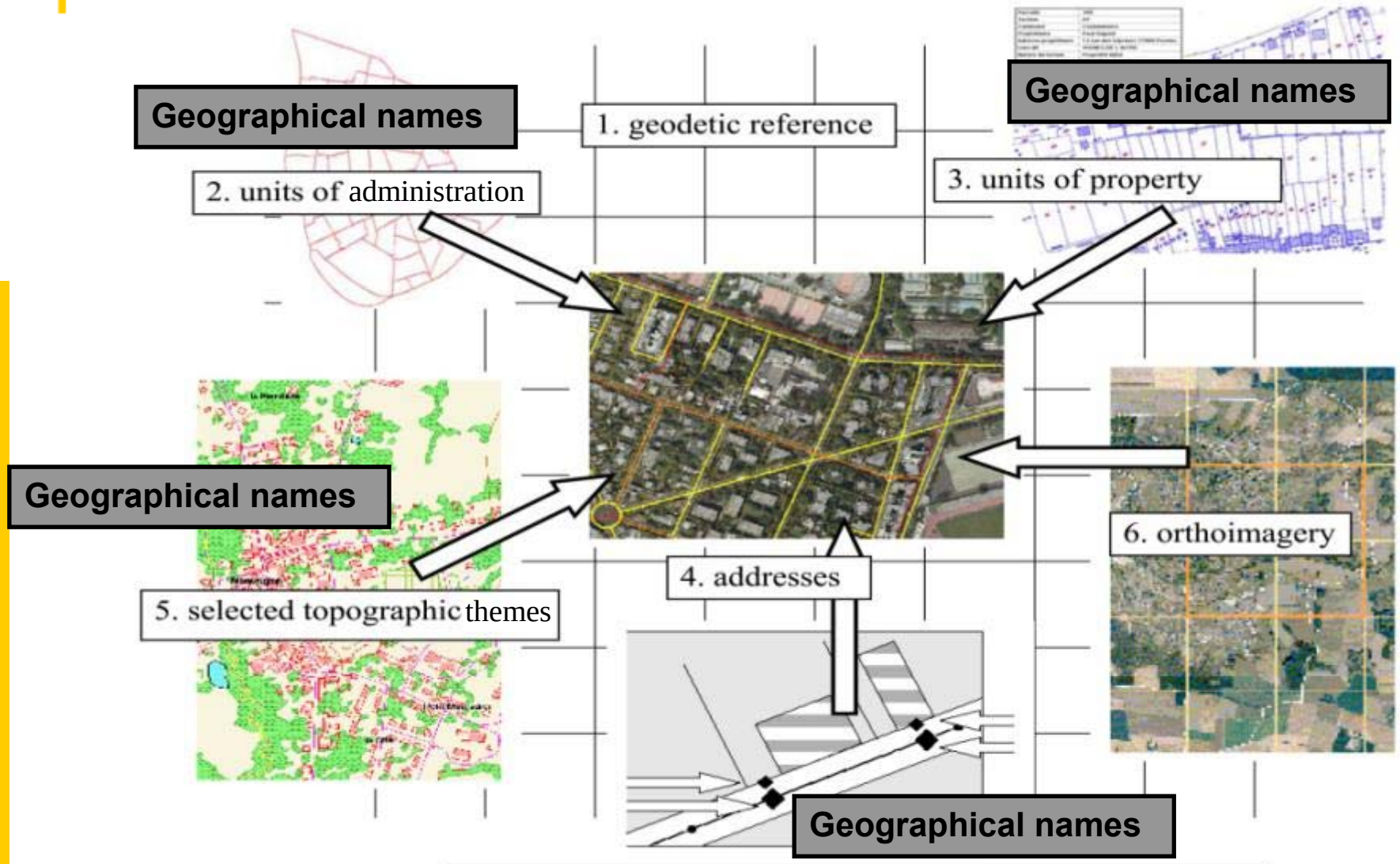
Information that is referenced to the earth‘s surface,
whether by coordinates (direct referencing) or by identifiers
such as addresses or postal codes or geographical names (indirect referencing).

Geo(graphic) data / spatial data:

Computer-readable geo information

Vector data model (feature data)	*Raster data model* (coverage data)
points, lines and polygons (areas)	gridded data (scanned maps, satellite images, orthophotos)

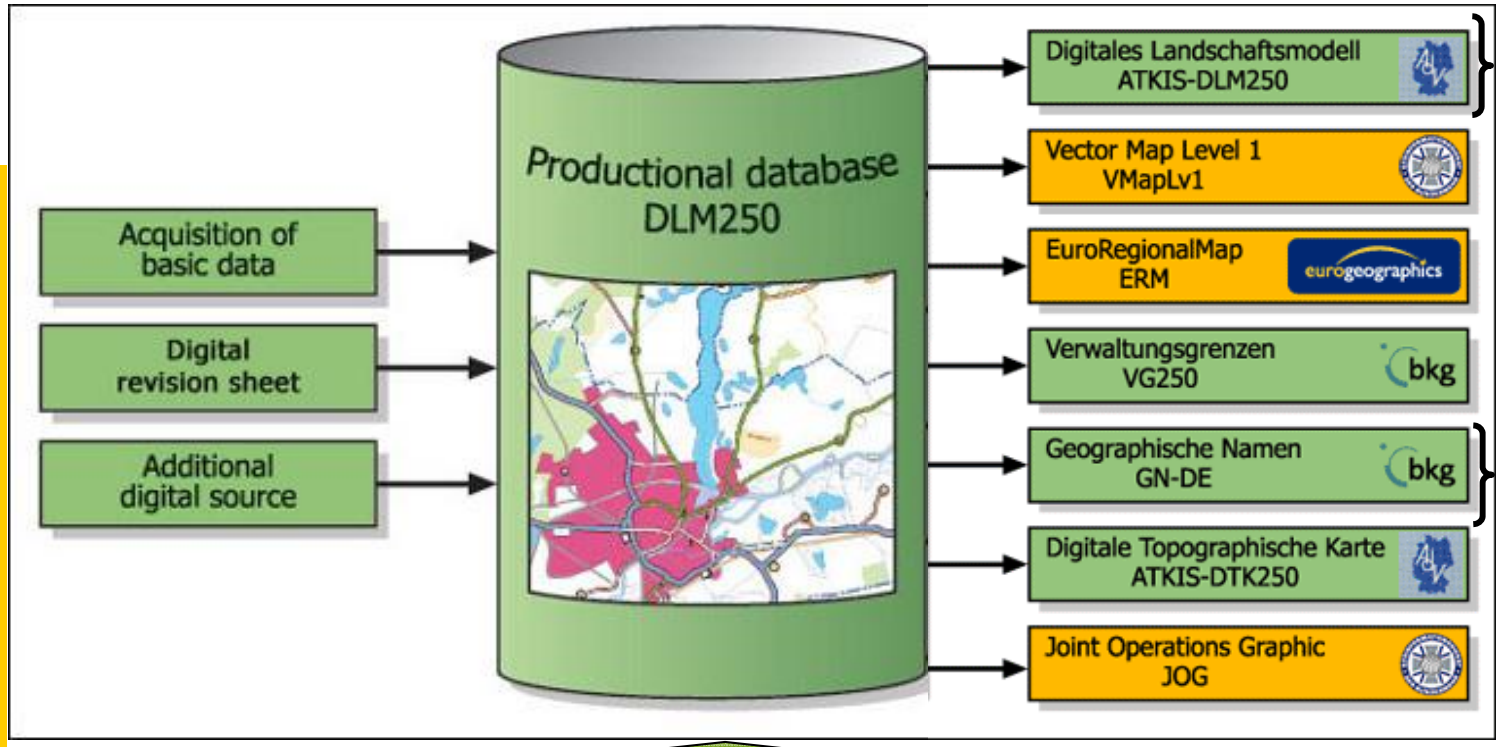
Names database and Geographic Information System (GIS)



The five components integrated in the sixth (common geodetic reference system) to create the complete set of

Names database and Geographic Information System (GIS)

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)**

Names databases and Web

Gazetteer

(printed)

Web

Names DATABASE

national

GIS

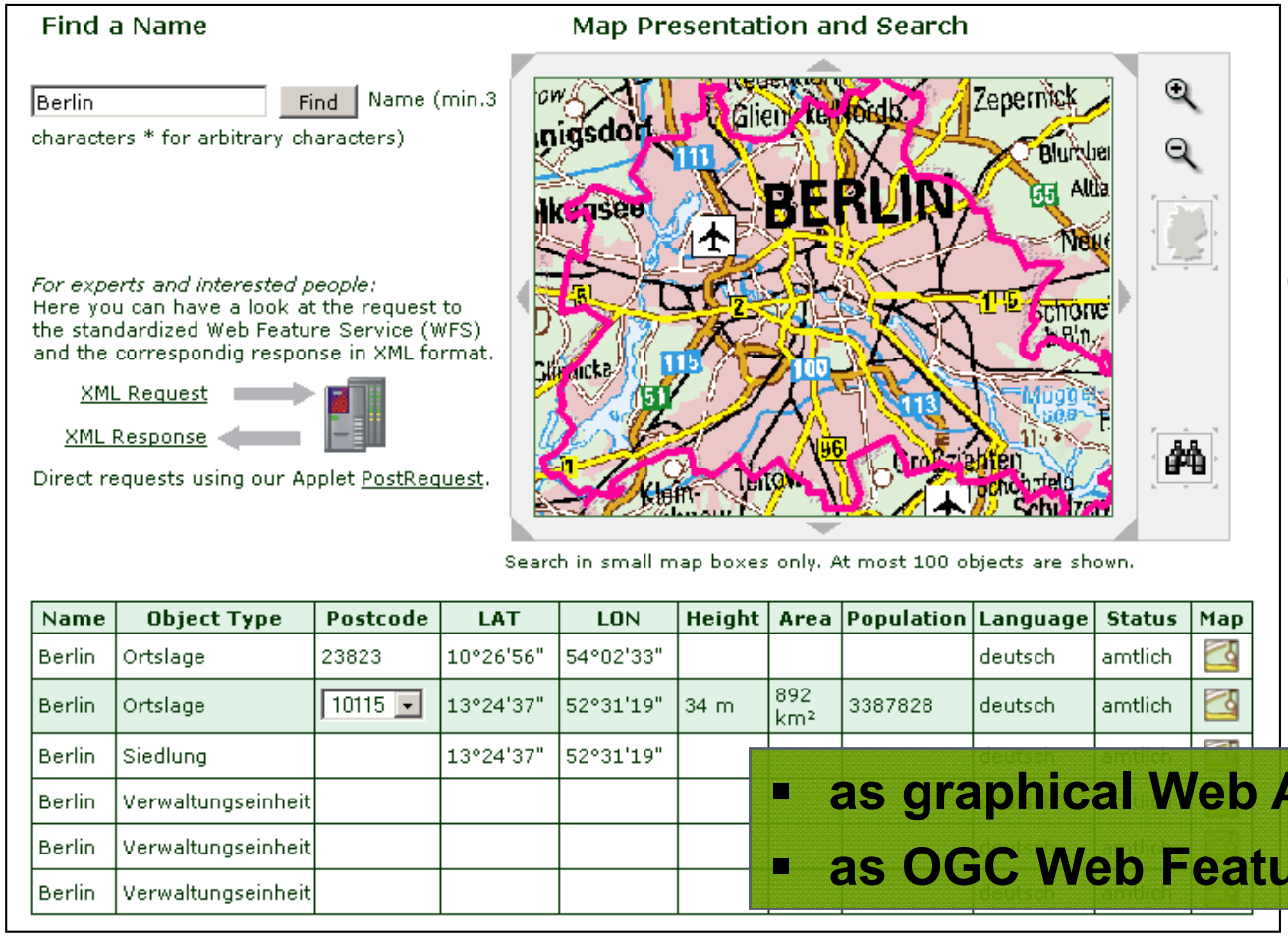
Map

Names DATABASE

international

The national gazetteer service –

search for geographical names of Germany…



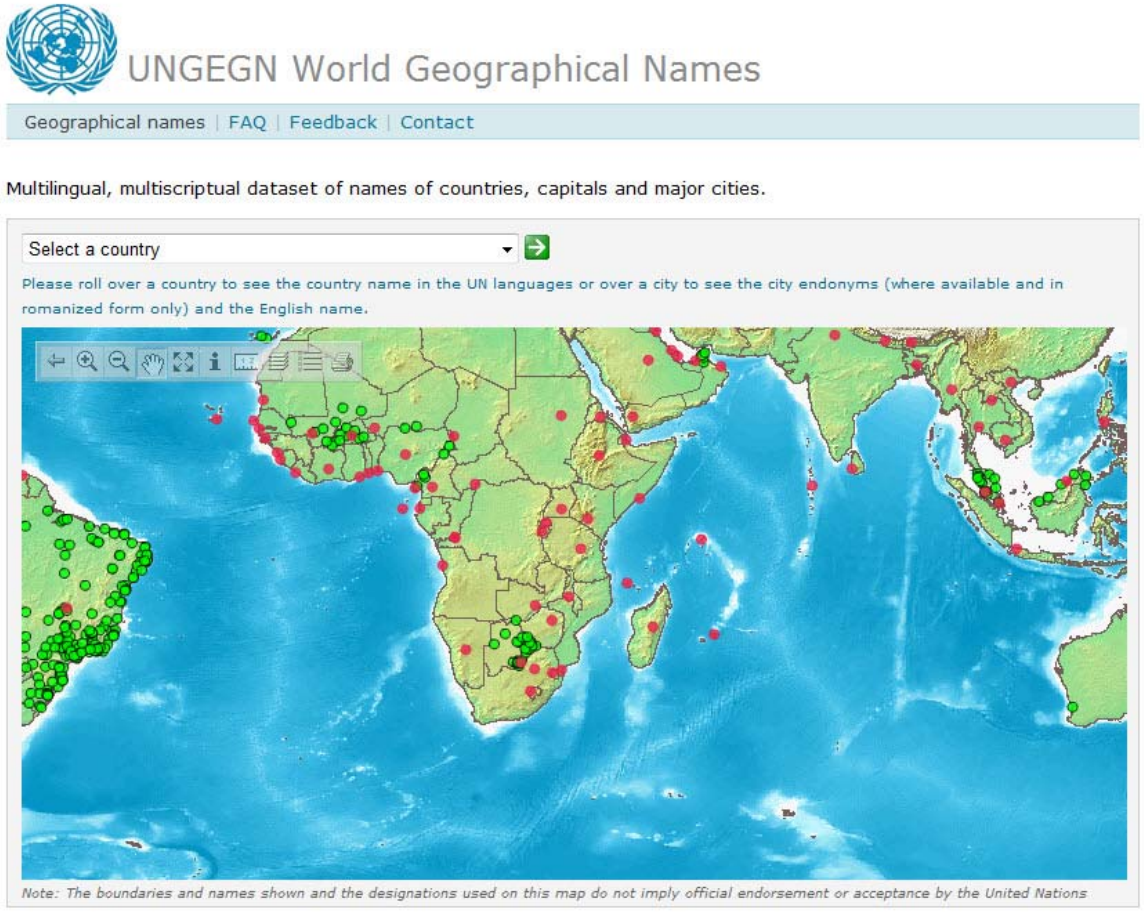
Name	Object Type	Postcode	LAT	LON	Height	Area	Population	Language	Status	Map
Berlin	Ortslage	23823	10°26'56"	54°02'33"				deutsch	amtlich	
Berlin	Ortslage	10115	13°24'37"	52°31'19"	34 m	892 km²	3387828	deutsch	amtlich	
Berlin	Siedlung		13°24'37"	52°31'19"						
Berlin	Verwaltungseinheit									
Berlin	Verwaltungseinheit									
Berlin	Verwaltungseinheit									

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

- **as graphical Web Application**
- **as OGC Web Feature Service (WFS)**

The international UNGEGN database –

search for geographical names of countries and capitals



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

Names databases and Web

UNGEGN World Geographical Names

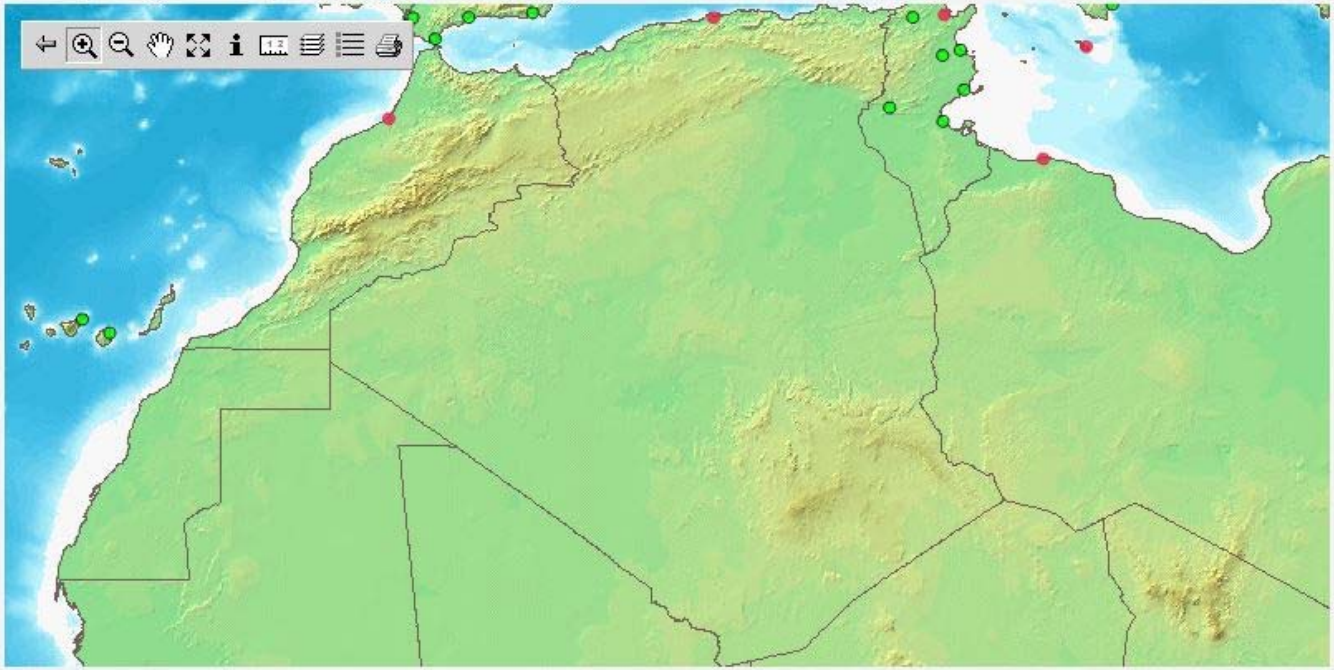
Multilingual, multiscriptual dataset of names of countries, capitals and major cities.

Geographical Names Database | About | FAQ | Feedback | Contact Beta version

Algeria

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.



Notes:

- The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the United Nations. The source for the file is UNGIWG. Final boundary between the Republic of Sudan and the Republic of South Sudan has not yet been determined.

- The Arabic fonts are not connected at this stage because of technical problems.

Names databases and Web

Country names

Endonyms

Language	Short name	Formal name
Arabic	الجزائر *Al Jazā'ir*	الجمهورية الجزائرية الديمقراطية الشعبية *Al Jumhūrīyah al Jazā'irīyah ad Dīmuqrāṭīyah ash Sha'bīyah*

United Nations languages

Language	Short name	Formal name
Arabic	الجزائر *Al Jazā'ir*	الجمهورية الجزائرية الديمقراطية الشعبية *Al Jumhūrīyah al Jazā'irīyah ad Dīmuqrāṭīyah ash Sha'bīyah*
Chinese	阿尔及利亚	阿尔及利亚民主人民共和国
English	Algeria	the People's Democratic Republic of Algeria
French	Algérie (l') [fém.]	la République algérienne démocratique et populaire
Russian	Алжир	Алжирская Народная Демократическая Республика
Spanish	Argelia	la República Argelina Democrática y Popular

Names databases and Web

Country names

Endonyms

Capital and other major cities

Capital city

Lat: 36.71 | Long: 3.08 الجزائر - Al Jazā'ir

Endonym

Language	City name	Source
Arabic	الجزائر *Al Jazā'ir*	UNGEGN

United Nations Languages

Language	City name	Source
Arabic	الجزائر	UN terms
Chinese	阿尔及尔	UN terms
English	Algiers	UN terms
French	Alger	UN terms
Russian	Алжир	UN terms
Spanish	Argel	UN terms

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
- Example: UNECA initiative – GeoNyms desktop database application
- Data model vs. database modelling

Use cases for Web applications using names data services

International purchase of properties
with Internet property services

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

Use cases for Web applications using names data services

Provision of European news

⇒ GN databases or web services can provide additional location based information

What are the purposes of a DB?

Gazetteer

(printed)

Web

Names DATABASE

national

GIS

Map

Names DATABASE

international

What are the purposes of a DB?

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.

What are the purposes of a DB?

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

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
- Example: UNECA initiative – GeoNyms desktop database application
- Data model vs. database modelling

Characteristics of databases

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

Different computer support different characteristics…

Characteristics of databases

Data Capture

↓

Data Storage

↓

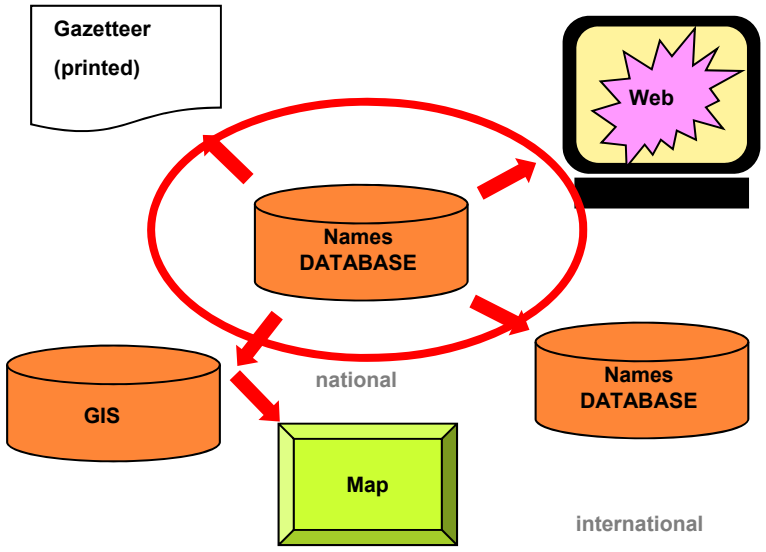
Data Processing

↓

Data Output

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

in a database



1. Operating system

An operating system (OS) is a software program that manages the hardware and software resources of a computer. 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 records stored in a computer in a systematic way, so that a computer program 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)

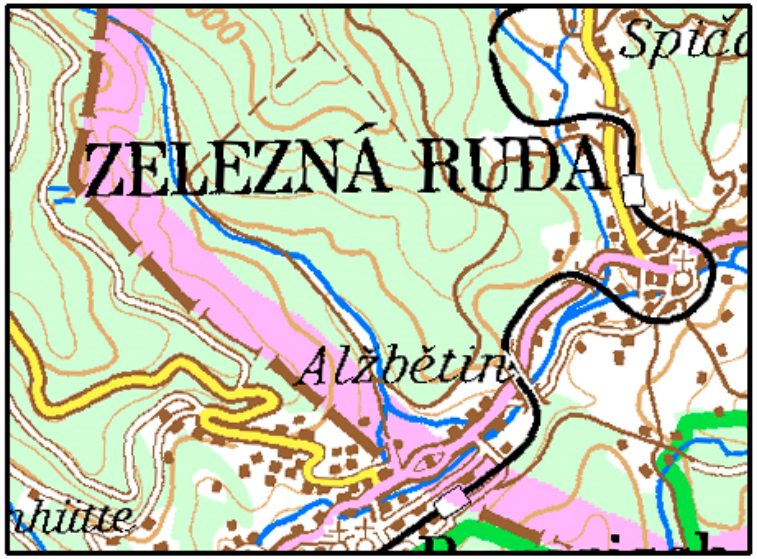
e.g. Codepages!!!

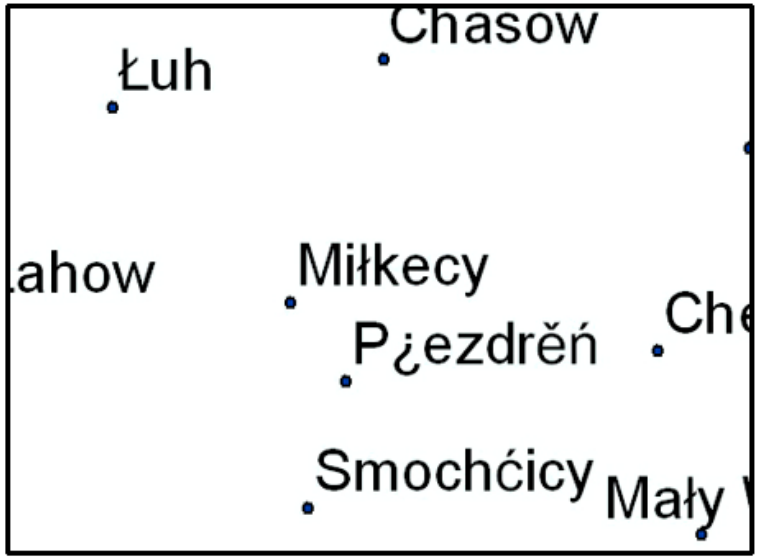
ISO8859 character set or ***Unicode?***
in (Microsoft Windows) operating system

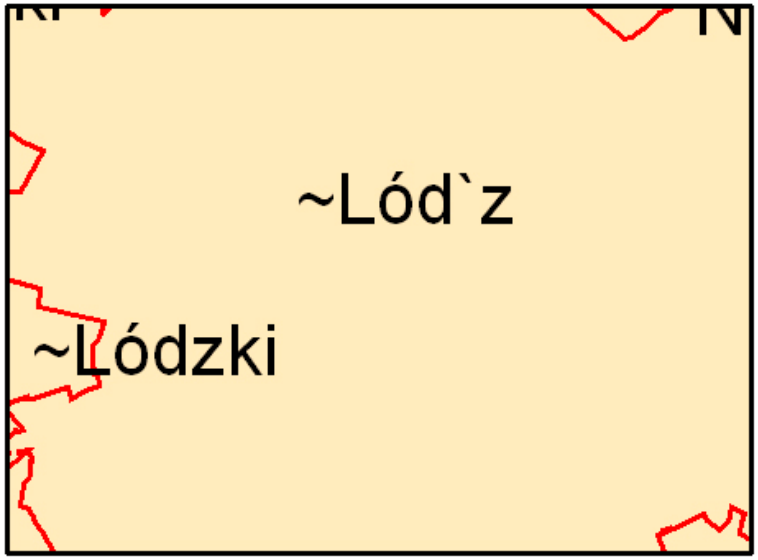
Different interpretation of characters possible in

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

Characteristics of databases







Operating System:

Microsoft Windows?

Linux? (open source)

Database:

Microsoft Access?

MySQL? PostGreSQL/PostGIS? (open source)

UNECA GeoNyms application? (open source)

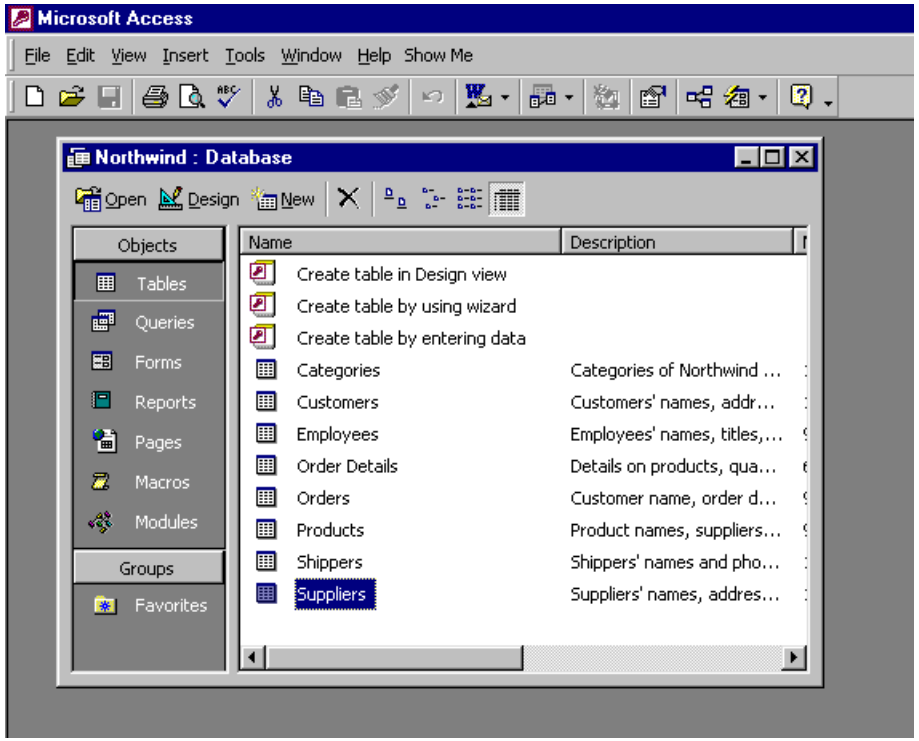
Geographical Information System

ESRI ArcGIS?

Quantum GIS? (open source)

Database:
Microsoft Access? (proprietary)

- Microsoft Access is a **pseudo-relational 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/



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/

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.

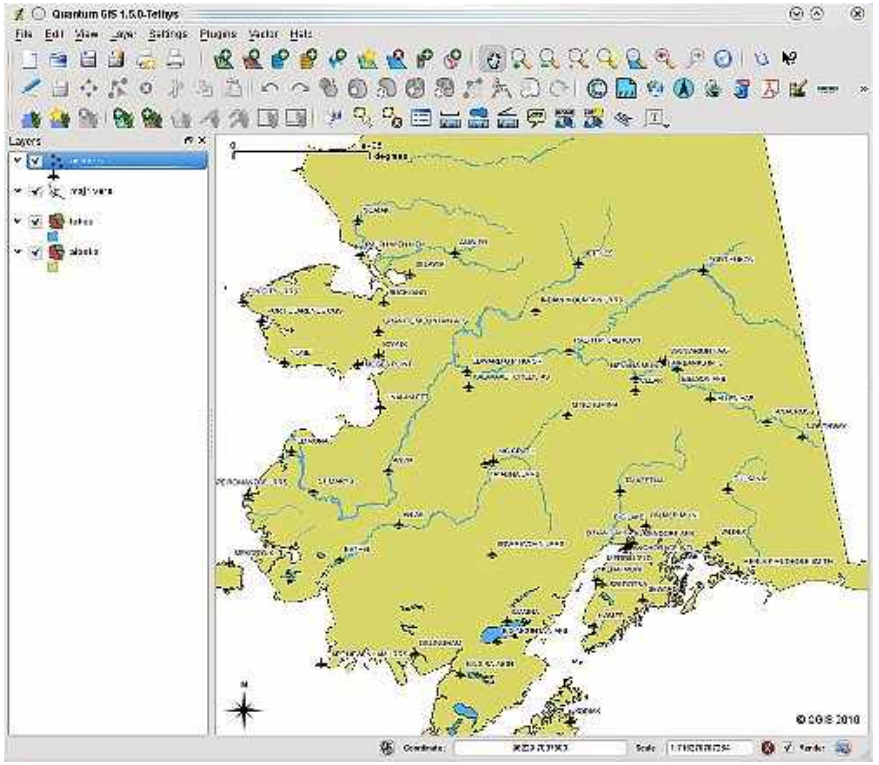
- ArcGIS **support and educational services** consist of technical maintenance programs, software releases and updates, technical support, online support services, publications training, and consulting services.
- Website: http://www.esri.com/software/arcgis/index.html

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

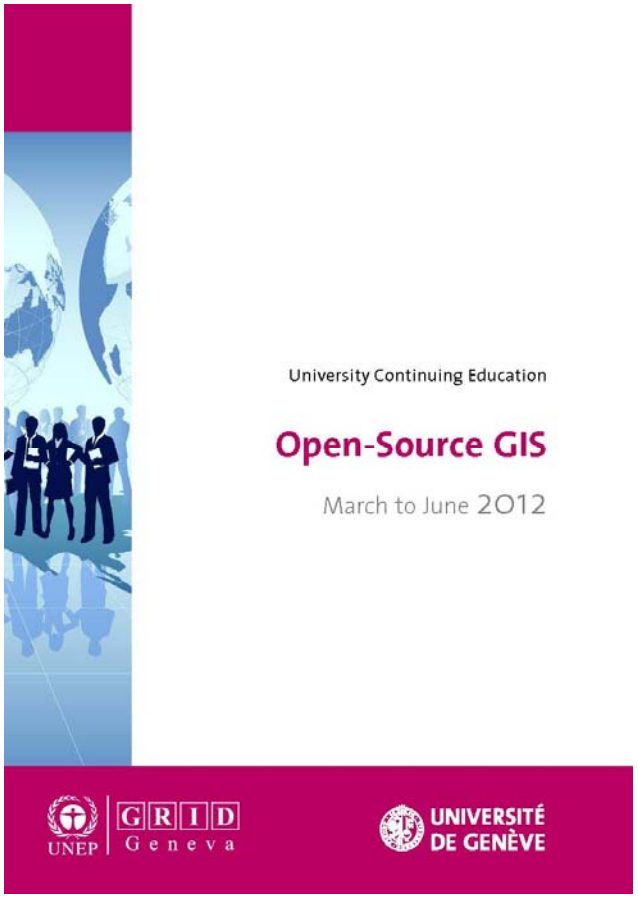
about:
OSGeo Live provides bootable ISO-Images and Virtual Machines which allow users to try out fully-operational versions of popular Free Geospatial Software without the need to install a thing.

Open your Mind through GIS: training in Open Source GIS

The University of Geneva in collaboration with UNEP/GRID-Geneva organizes *continuing education training on Open Source GIS*. It aims at presenting and discussing major open-source desktop GIS solutions.

The next one will be organized April - May 2015 (Inscription: 28 February 2015)

More information on:
http://www.unige.ch/formcont/opengis.html



Content

Preparing the names database – General issues

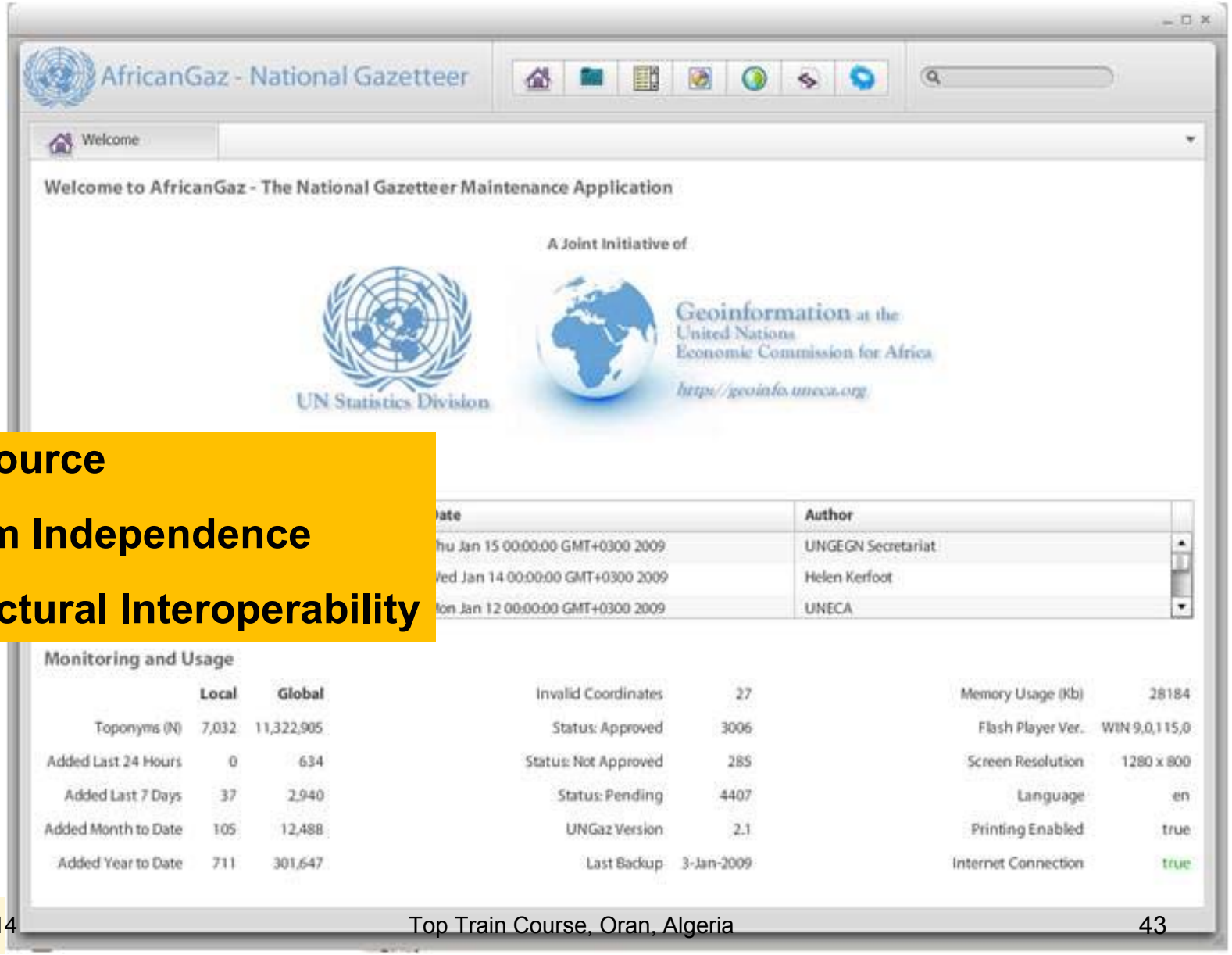
- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names database and Web
- Use cases for Web applications

Description of database characteristics – Database management

- Characteristics of databases
- Realisation of databases: open source vs. proprietary
- Example: UNECA initiative – GeoNyms desktop database application
- Data model vs. database modelling

Federal Agency for Cartography and Geodesy

UNECA initiative – GeoNyms desktop application



- **Open source**
- **Platform Independence**
- **Architectural Interoperability**

UNECA initiative – GeoNyms desktop application

Multilingual

English and French are implemented … Arabic coming

Adaptable to National settings

1. Feature Types
2. Feature Statuses
3. Administrative Divisions

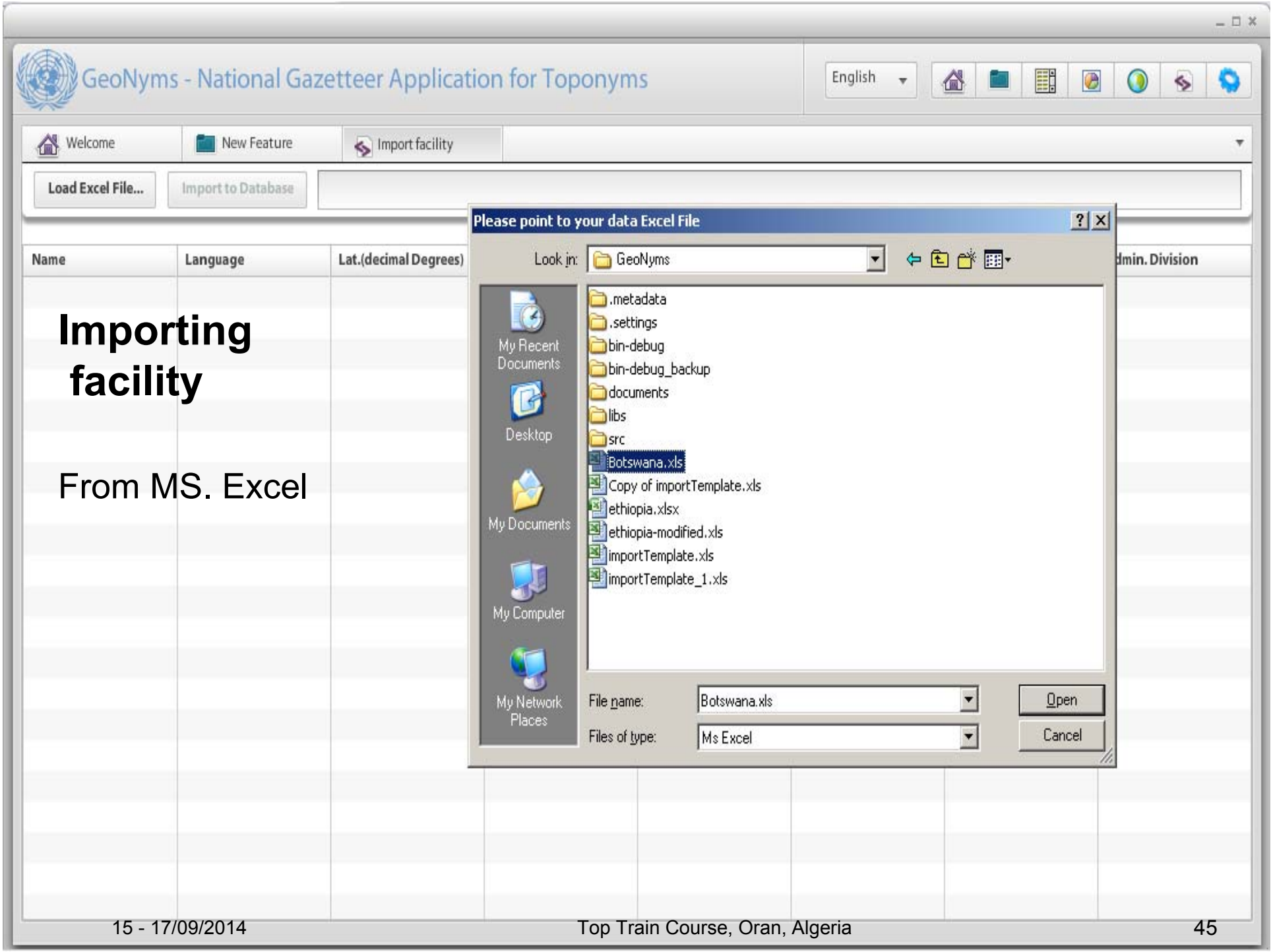
Filtering & Exporting facility

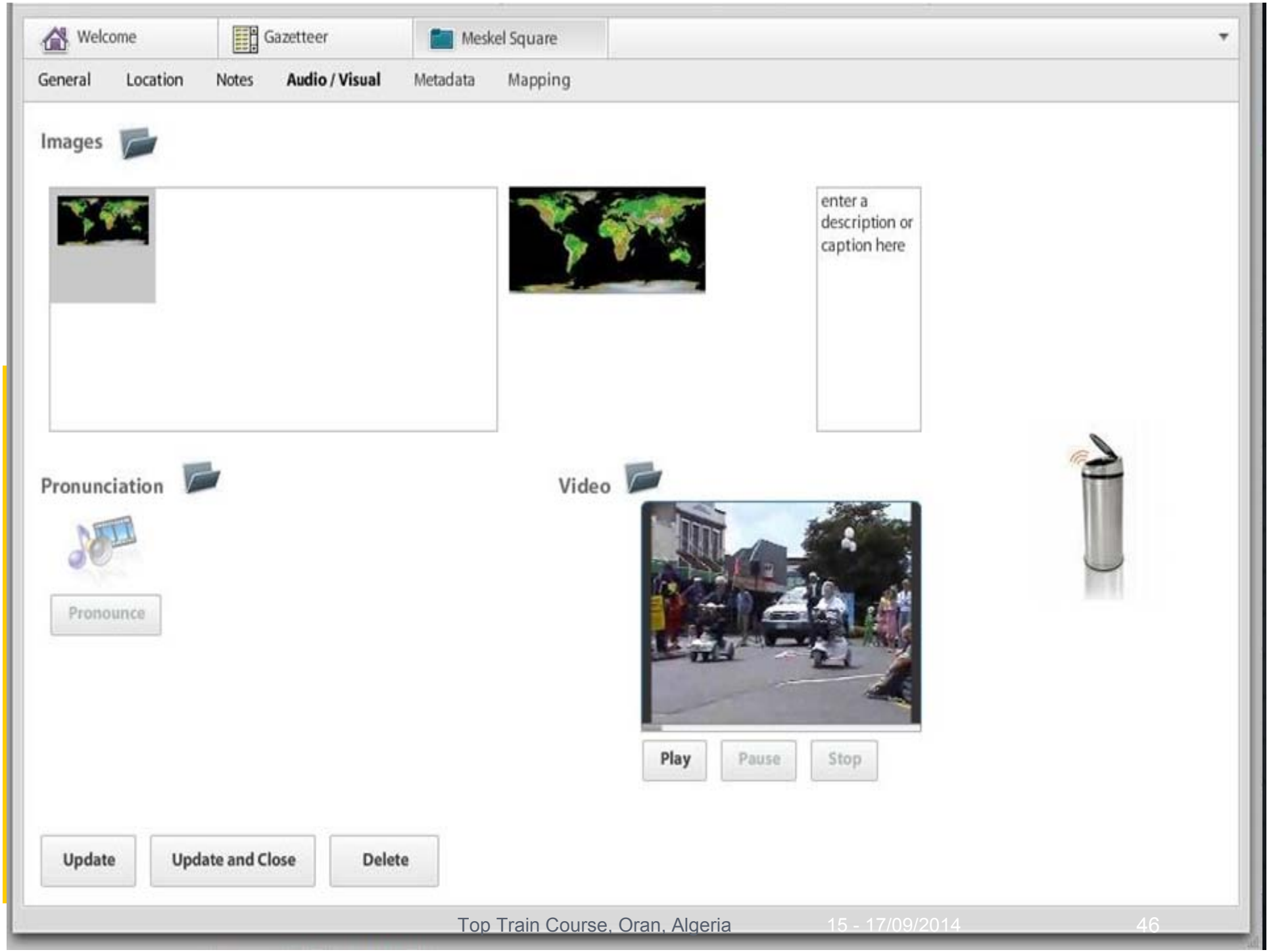
Implemented now

1. *.xls (Excel)
2. *.csv (Textfile)

Planned:

3. Shape files
4. XML


GeoNyms - National Gazetteer Application for Toponyms
English
Welcome
New Feature
Import facility
Load Excel File...
Import to Database
Name
Language
Lat.(decimal Degrees)
dmin. Division
Importing facility
From MS. Excel
Please point to your data Excel File
Look in:
GeoNyms
.metadata
.settings
bin-debug
bin-debug_backup
documents
libs
src
Botswana.xls
Copy of importTemplate.xls
ethiopia.xlsx
ethiopia-modified.xls
importTemplate.xls
importTemplate_1.xls
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
File name:
Botswana.xls
Files of type:
Ms Excel
Open
Cancel

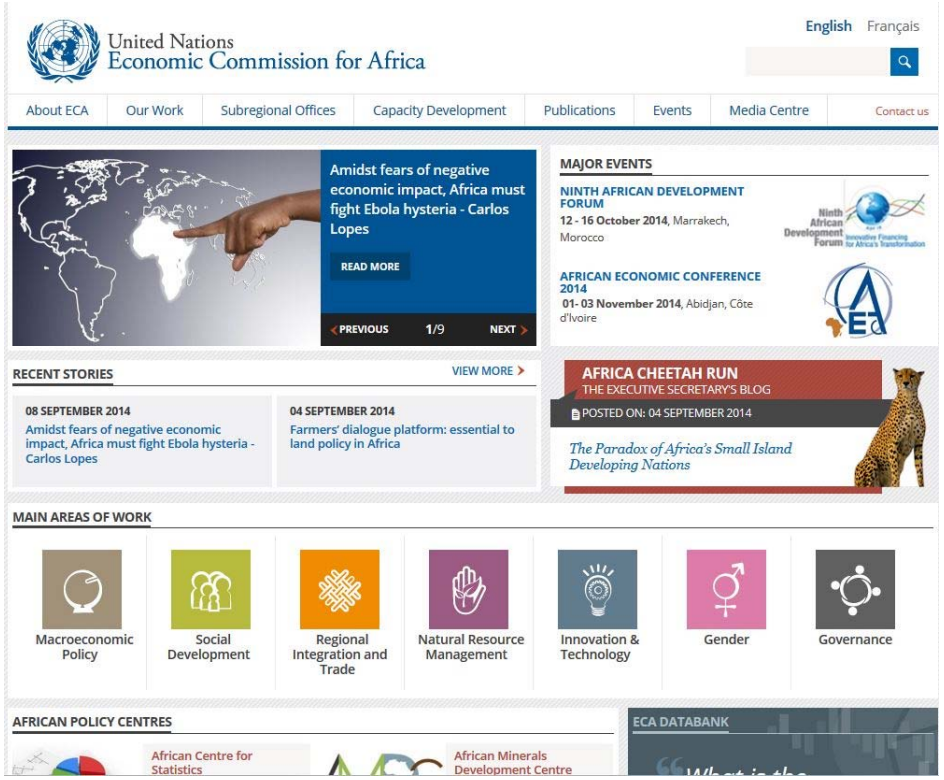

Welcome
Gazetteer
Meskel Square
General
Location
Notes
Audio / Visual
Metadata
Mapping
Images
enter a description or caption here
Pronunciation
Pronounce
Video
Play
Pause
Stop
Update
Update and Close
Delete

Federal Agency for Cartography and Geodesy

UNECA initiative – GeoNyms desktop application

How to get it?

Downloads/documentation in
http://geoinfo.uneca.org/geonyms



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
- Example: UNECA initiative – GeoNyms desktop database application
- Data model vs. database modelling

Data model – 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 unique identifier 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

Data model – database modelling

The European (INSPIRE) GN schema: elements

mandatory

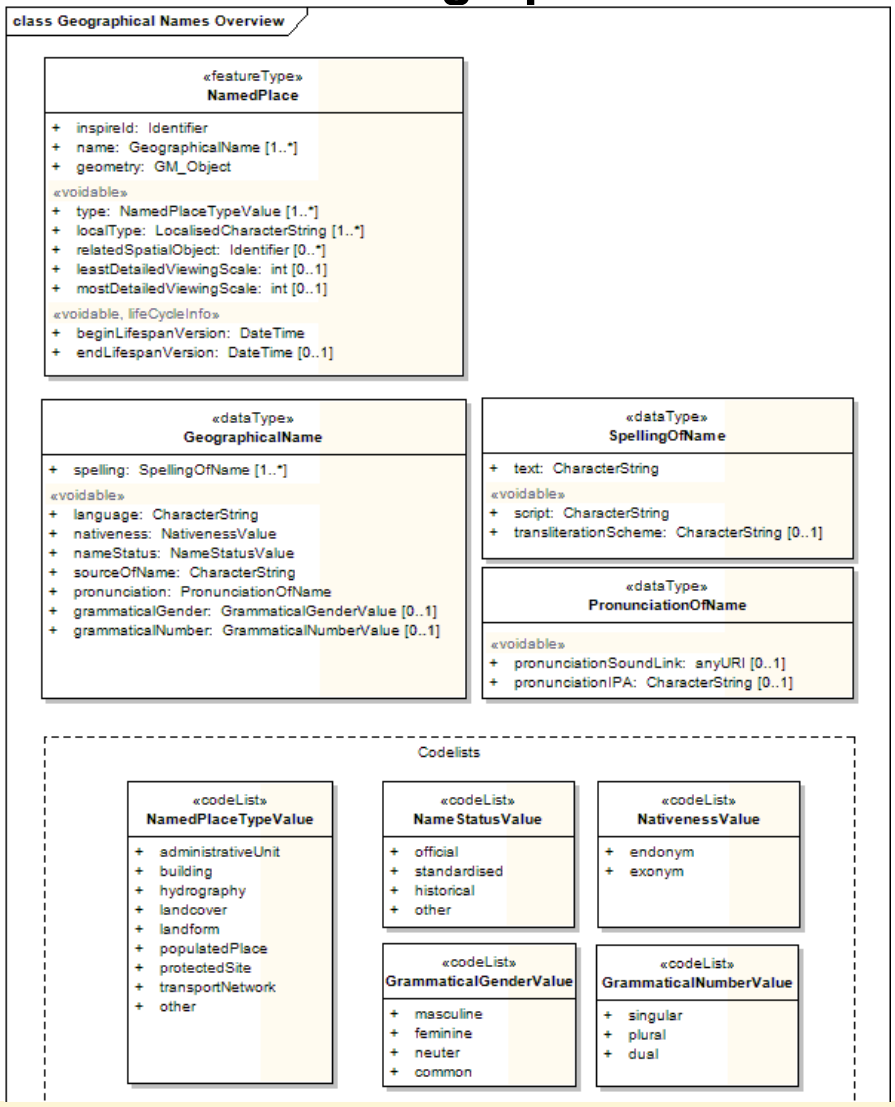
- name(s) (text, spelling)
- geomety
- feature type
- unique identifier

'voidable'

- 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)
- …

Data model – database modelling

INSPIRE Geographical Names – UML Schema



A <u>Named Place</u>,
representing a real world entity referred to by a Geographical Name

e.g "the City of Athens" type = built-up area
geometry = {X, Y}

is associated with one or several <u>Geographical Names</u>,
i.e. proper noun applied to the feature

(1) "Athína" language = Greek
nativeValue = Endonym

(2) "Athens" language = English
nativeValue = Exonym

and may have one or several <u>Spellings of Name.</u>
i.e. proper way of writing the name

(1.1) text = **Aθnva**
script = Greek

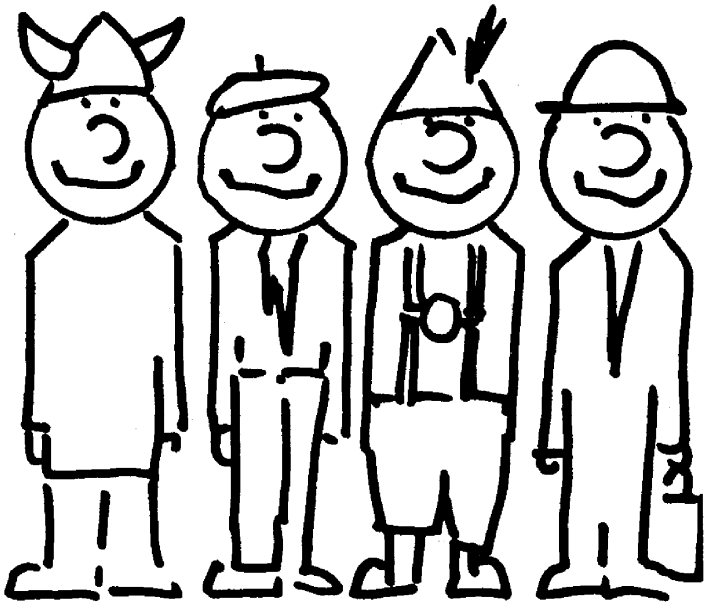
(1.2) text = **Athína**
script = Latin

(2.1) text = **Athens**
script = Latin

Standardization in Europe

Not standardized and very different

Almost harmonized, but with national specialities



Thank you for your attention!