### INTERNATIONAL TRAINING ON TOPONYMY

# MODULES

DAY 5

19 - 23 JUNE 2023

BALI, INDONESIA

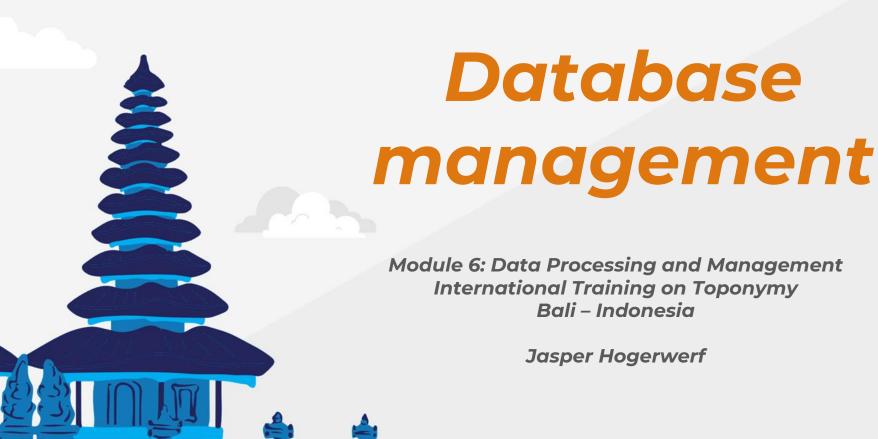


United Nations
Group of Experts on











#### Content

#### **Preparing the names database – General issues**

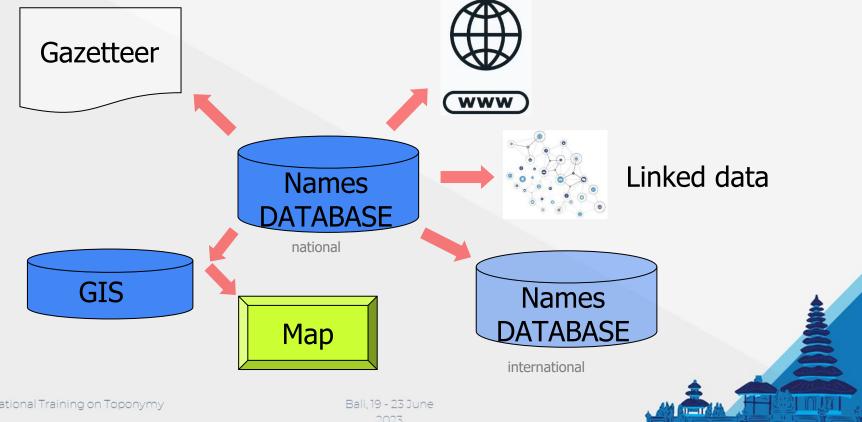
- What are the purposes of a database?
- Names database and Geographic Information System (GIS)
- Names databases and the Web
- Use cases for Web applications

#### **Description of database characteristics – Database management**

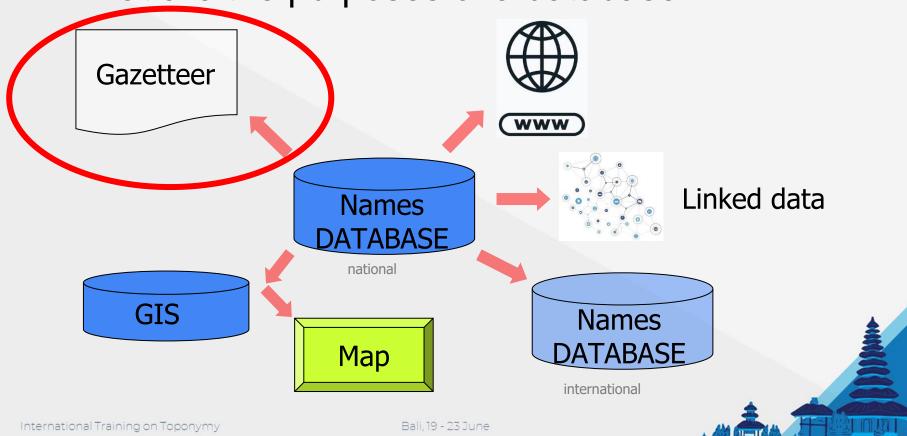
- Characteristics of databases
- Realization of databases: open source vs. proprietary
- Data model and database modelling









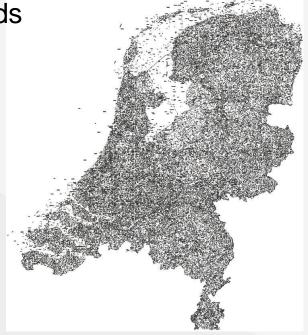








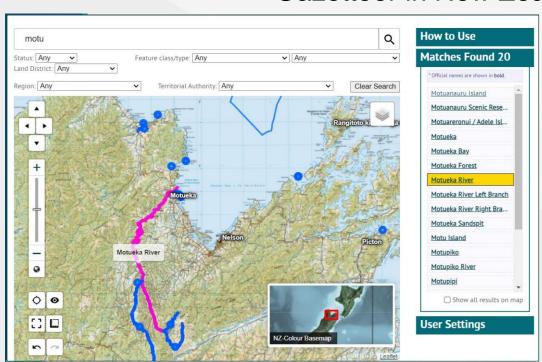
20 years ago: Digital military gazetteer Topografische Dienst



Current: TOPnamen annotation gazetteer Kadaster



#### Gazetteer in New Zealand



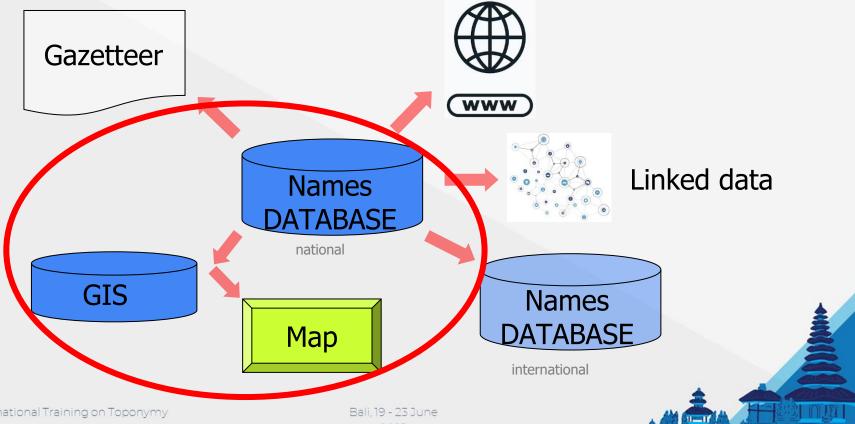
Gazetteer web map application

Toitū Te Whenua Land Information New Zeeland

https://gazetteer.linz.govt.nz/















- 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

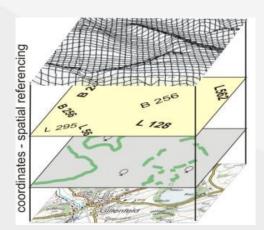


#### **Geographic Information System (GIS)**

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)

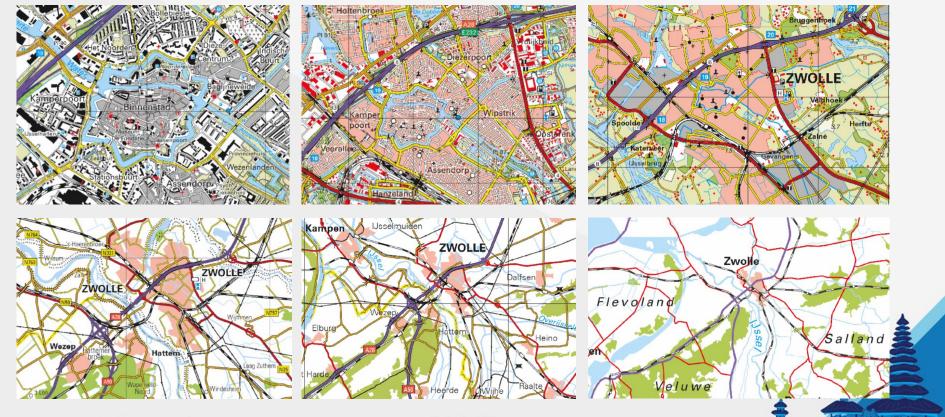




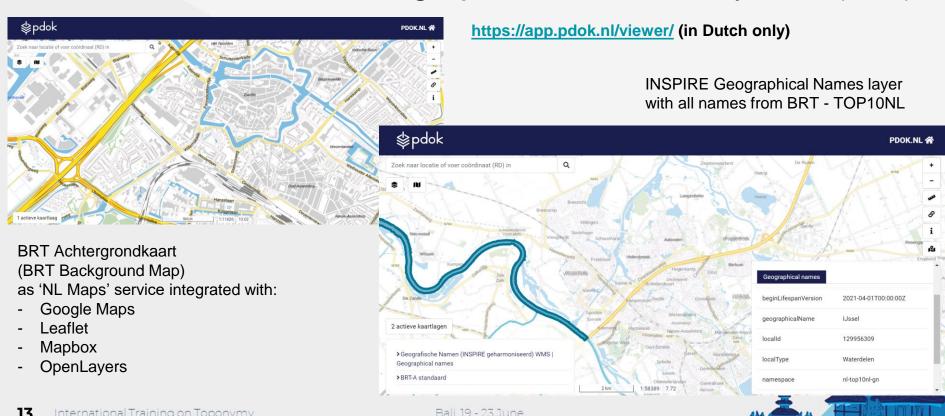




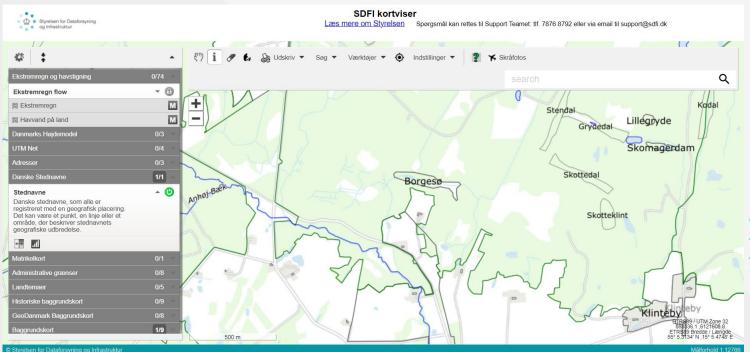












Example of a separate names database / register:

**Danske Stednavne** 

Styrelsen for Dataforsyning og Infrastruktur, Denmark

https://sdfi.dk/data-omdanmark/voresdata/danske-stednavne (in Danish)



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.





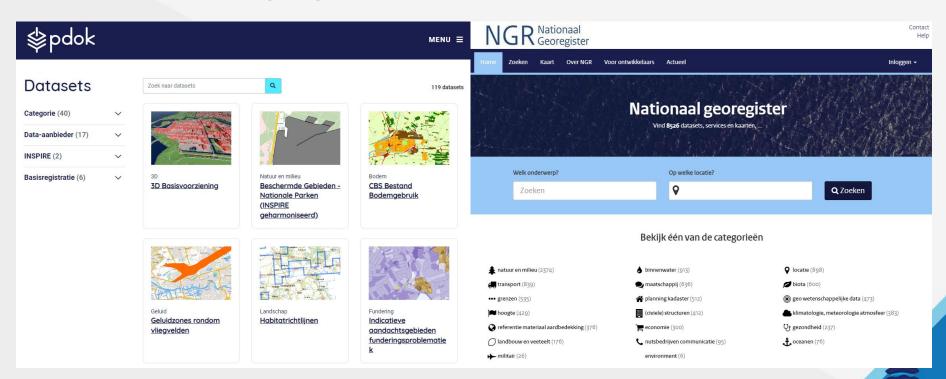
A spatial data infrastructure (SDI) is a data infrastructure implementing a framework of geospatial/geographic data, metadata, users and tools that are interactively connected in order to use spatial data in an efficient and flexible way.

#### 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

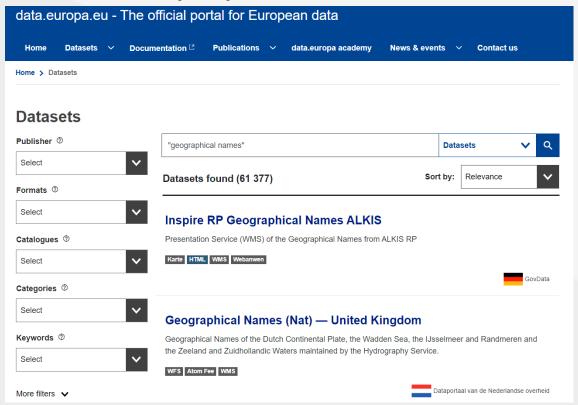




https://www.pdok.nl/ (in Dutch)

https://www.nationaalgeoregister.nl/ (in Dutch)

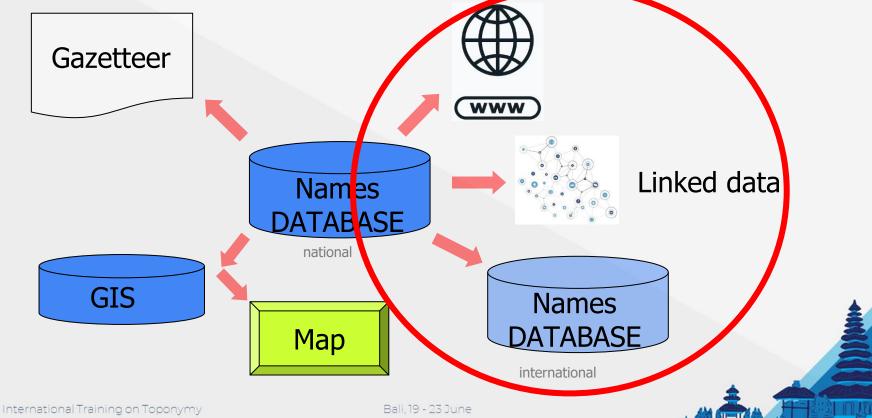




https://data.europa.eu/en

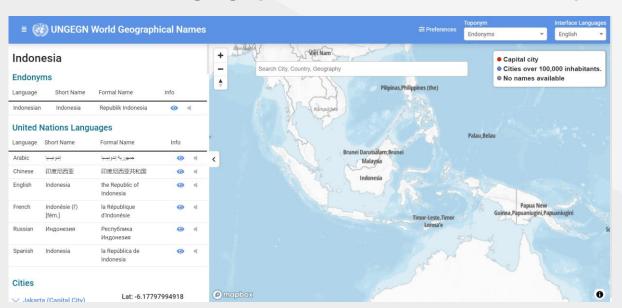








# UNGEGN World Geographical Names database search for geographical names of countries and capitals

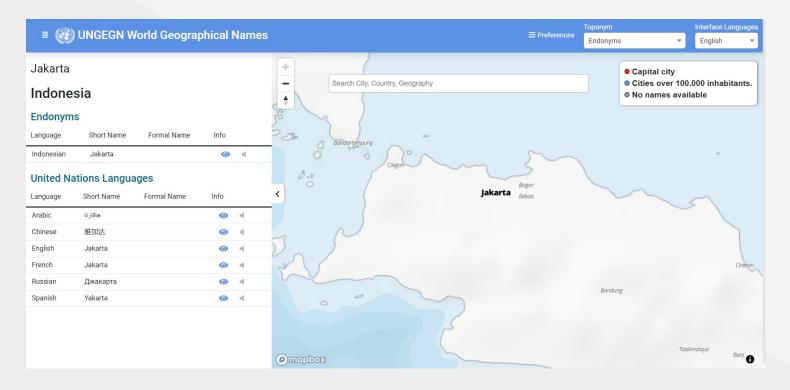


https://ungegn-dashboardhja6b0dafqh0c6dd.z01.azurefd.net/

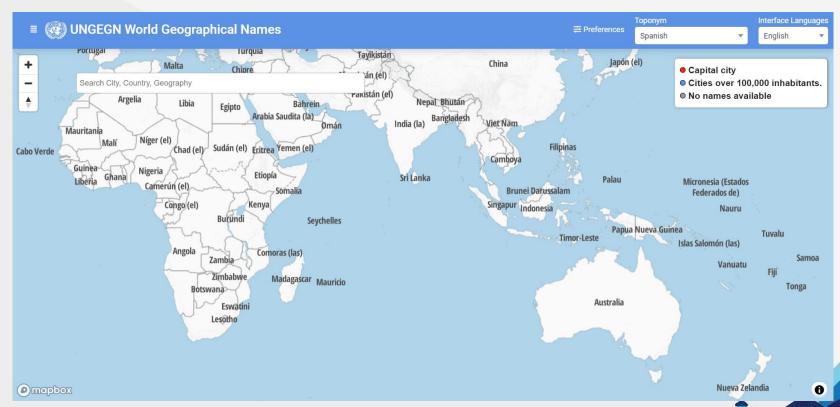
http://unstats.un.org/unsd/geo info/geonames/







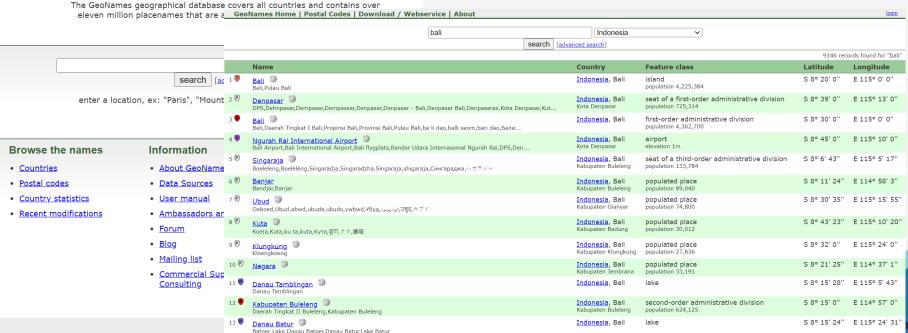




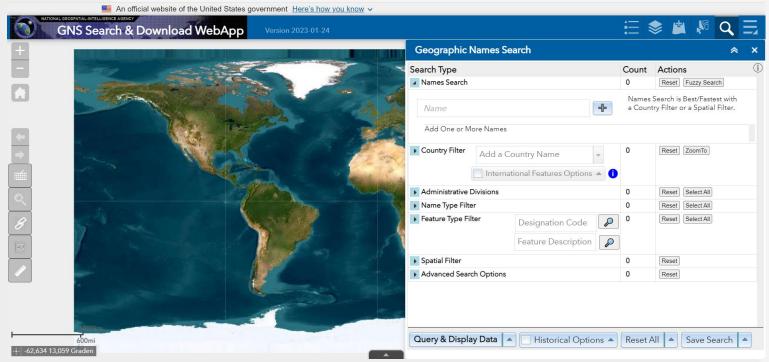




#### https://www.geonames.org/



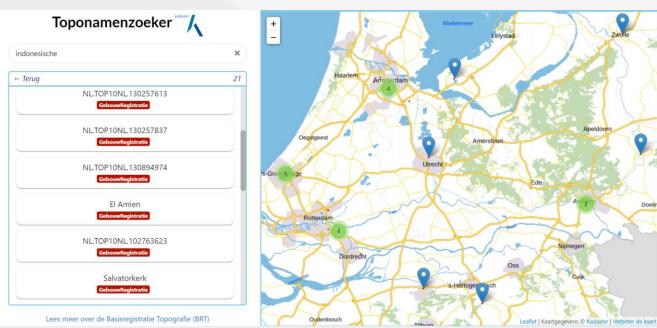




https://geonames.nga.mil/geonames/GeographicNamesSearch/



#### **Toponamenzoeker (Topo Names Finder)**

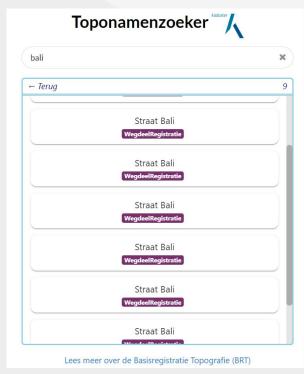


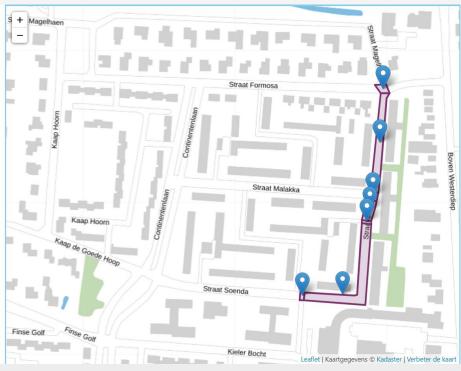
Web application to search for geographical names in the Key Register Topography (BRT) of the Netherlands















**Developing towards linked open data:** 

1 star: Map with names online

2 stars: Names database available online in Esri format (e.g. File Geodatabase)

3 stars: Names database available online in open source format (e.g. Geopackage)

4 stars: Names, attributes and values available as URI (linkable web data)

5 stars: URIs of names database linked to other linked open data



#### Content

#### **Preparing the names database – General issues**

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

#### **Description of database characteristics – Database management**

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





### Use cases for Web applications using names data



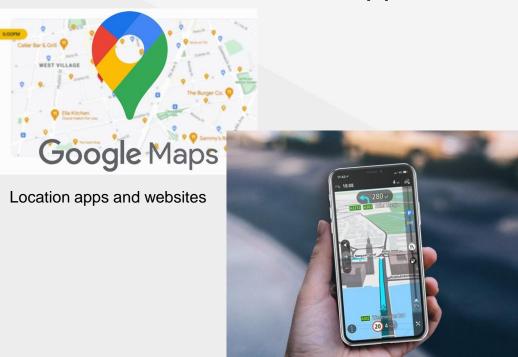
International purchase of properties with Internet property services <a href="http://www.viviun.com/">http://www.viviun.com/</a>

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



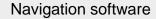


### Use cases for Web applications using names data



Select a shipping address Add a new address Or search for an Amazon Pickup location near you Country/Region Amazon Pickup locations enable you to pick up your package at your convenience United States Search by Address Full name (First and Last name) e.g.: 333 Boren Ave N, Seattle John Nawma Search by Zip Code: e.g. 98109 Search by Landmark: +16506305555 e.g. Space Needle Search by Locker or Store Street address or P.O. Box e.g.: Juno Apt, suite, unit, building, floor, etc Make this my default address Delivery instructions (ontional) Use this address

Web shops using address data





#### Content

#### **Preparing the names database – General issues**

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

#### **Description of database characteristics – Database management**

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







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





Data Capture



Data Storage

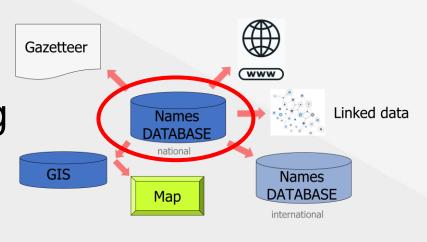


**Data Processing** 



**Data Output** 

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





#### Main software requirements:

#### 1. Operating system

software program that manages the hardware and software resources of a computer.

<u>Examples</u>: Microsoft Windows, Apple iOS (proprietary), Linux (open source = free of charge and usage)

#### 2. Database

collection of records stored in a computer in a systematic way, so that it can be consulted to answer questions.

Examples: Microsoft Access, SQL Server, Oracle (proprietary), MySQL, PostGreSQL / PostGIS (open source)

#### 3. Geographic Information System (GIS) software

tool that allows users to select, analyze, edit, manipulate and visualize data from a database.

Examples: ESRI ArcGIS, Hexagon Geomedia, MapInfo (proprietary), QGIS, GRASS GIS (open source)

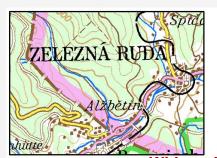


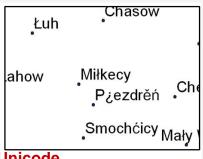
### **Important: Code files!**

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













#### Content

#### **Preparing the names database – General issues**

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

#### **Description of database characteristics – Database management**

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





## **Operating System:**

Microsoft Windows?

Linux? (open source)

#### Database:

Microsoft Access?

MySQL? (open source)

PostgreSQL/PostGIS? (open source)

## **Geographical Information System**

**ESRI ArcGIS?** 

Quantum GIS? (open source)

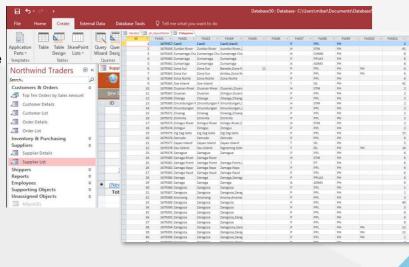




## **Database:**

## 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: <a href="https://www.microsoft.com/en-us/microsoft-365/access">https://www.microsoft.com/en-us/microsoft-365/access</a>





## Database:

MySQL? (open source)

#### "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:

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: <a href="http://www.postgresql.org/">http://www.postgresql.org/</a>

http://www.postgis.net/ - http://postgis.refractions.net/



**Geographical Information System** 

ESRI ArcGIS? (proprietary)

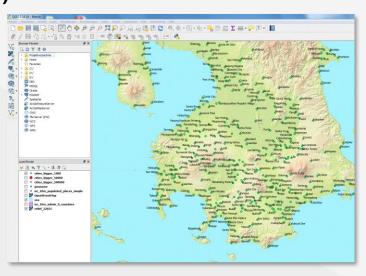
- Esri® ArcGIS® facilitates collaboration and lets you author data, maps, globes, and models on the desktop and serve them for use on a desktop, in a browser, or in the field, depending on the needs of your organization.
- ArcGIS support and educational services consist
  of technical maintenance programs, software
  releases and updates, technical support, online
  support services, publications, training, and
  consulting services.
- Website: <a href="https://www.esri.com/en-us/arcgis/about-arcgis/overview">https://www.esri.com/en-us/arcgis/about-arcgis/overview</a>





# Realization of databases Geographical Information System Quantum GIS? (open source)

- 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/site





# **Open Source Software Collection**

Open source software collection: OSGeo Live

URL: <a href="http://live.osgeo.org/en/index.html">http://live.osgeo.org/en/index.html</a>

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.



#### Content

#### **Preparing the names database – General issues**

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

#### Description of database characteristics – Database management

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





# Data model and 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



# Data model and database modelling

#### The European (INSPIRE) GN schema: elements

#### mandatory

'voidable'

- 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)

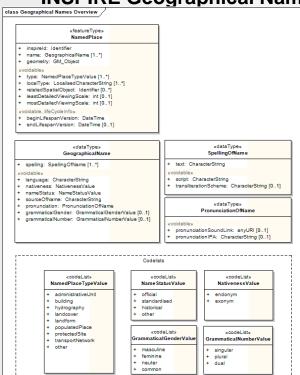
...





# Data model and database modelling

#### **INSPIRE Geographical Names – UML Schema**



#### A Named Place,

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 **Spellings of Name.**

i.e. proper way of writing the name

(1.1) text = 
$$\mathbf{A}\theta\mathbf{n}\mathbf{v}\mathbf{a}$$
 (2.1) text =  $\mathbf{A}\mathbf{t}\mathbf{h}\mathbf{e}\mathbf{n}\mathbf{s}$  script =  $\mathbf{L}\mathbf{a}\mathbf{t}\mathbf{i}\mathbf{n}$  (1.2) text =  $\mathbf{A}\mathbf{t}\mathbf{h}\mathbf{i}\mathbf{n}\mathbf{a}$ 

(1.2) text = Atninascript = Latin



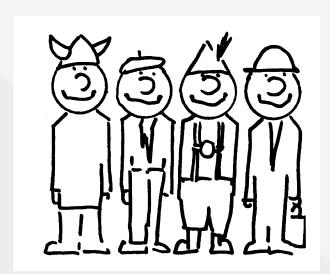
# Standardization in Europe

Not standardized and very different



© Andreas Illert

Almost harmonized, but with national specialities

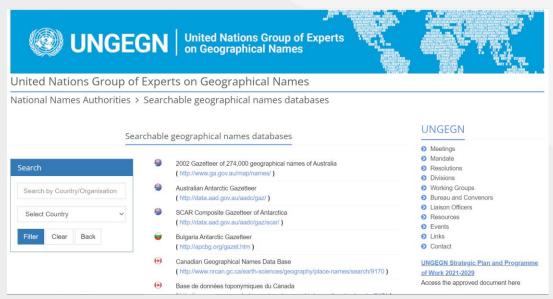


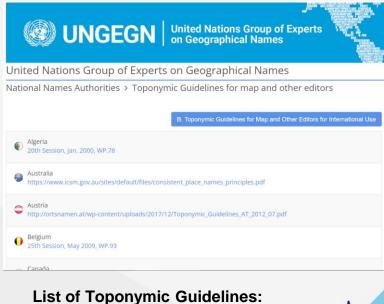


# Information on geographical names data

#### List of geographical names databases:

https://unstats.un.org/unsd/ungegn/nna/geo-names/





https://unstats.un.org/unsd/ungeg

n/nna/toponymic/



## Get involved in UNGEGN

### **Working Group on Geographical Names Data Management**



https://unstats.un.org/unsd/ungegn/working\_groups/wg2.cshtml





## Thank you for your attention!

