

First EuroGeoNames Workshop of the Consortium & Reference Group

Frankfurt am Main, 23. – 24.10.2006

**Presentation of the synopsis of the state
of the art of geographical names data
information from 15 Reference Group
member countries**



Synopsis was made by:

- Summarizing of detailed questionnaires and SI-EGN Final Report (for 15 countries)
- Checking reference documents (for 15 countries)
- Checking test data (for 14 countries)

Database scenarios:

- One digital stand-alone database – 9 (AT, CZ, DE, FI, HU, LV, SI, NL, NO)
- Distributed databases - 1 (FR)
- Integrated databases for different scales - 3 (CY, DE, SK)
- Mix between one digital stand-alone database and distributed databases – 2 (ES, TR)
- Mix between one digital stand alone database and integrated databases for different scales - 1 (LT).

Prevailing models are one digital stand-alone database and integrated databases for different scales.

Data of geographical names (at less some of them) of 11 countries (CZ, CY, DE, FI, FR, LT, NL, NO, SK, ES, TR) are linked to topographic or cartographic data.

Operating system:

- MS Windows is used in 13 countries
- LINUX in other two countries (FR and SK)
- UNIX is used together with MS Windows in 2 countries (CY, FI).

Database package:

- Oracle - 9 (AT, CY, CZ, FI, DE, NL, NO, SI, ES)
- Follow MS Access - 4 (DE, HU, ES, TR)
- SQL – 3 (DE, FR, LV)
- DBF – 1 (SK)
- LT has Geodatabase (national reference datasets) package.

Germany has data stored in 3 different database packages (Oracle, MS Access and SQL) and Spain in two (Oracle and MS Access).

Language codepages:

- ISO 8859 standard – all
- MS Windows Codepages (1250 – 1257) - 9 (CY, CZ, DE, HU, LV, LT, SI, ES, TR)
- MS-DOS Codepage - 6 (CY, CZ, DE, HU, LV, TR).
- Unicode - 4 countries (DE, LV, ES, SK)

Data model:

- Descriptions for some countries are not provided or are not detail.
- The principal difference among data models is that geographical names in some countries are part of topographic databases while in others names are stored in geographical names registers or databases.
- If geographical names are part of topographic databases (LT, NL) names are usually stored as an attribute value of a feature object of feature type. There is no special names data model but data model of topographic database.
- If names are stored in geographical names registers/databases data model contains two or more linked tables of entity types or of different codes.

Names are derived from the maps:

- Of middle scales – 11 countries:
 - 1 : 20 000 (FI)
 - 1 : 25 000 (FR, SK, SI, ES, NL, TR-populated places)
 - 1 : 50 000 (AT, HU, LV, LT, NO)
- Of large scale - 7 countries
 - 1 : 5000 – (CY, NO, SI)
 - 1 : 10 000 – (CZ, LT, SK, NL)
- Of small scales (1 : 250 000) – 4 countries (DE, SK, SI, TR-other than populated)

Feature categorisation:

- NATO standard - 3 countries (SK, NL, TR)
- National feature coding - 12 countries

National feature coding structures are divided into two types as:

- Feature structure of topographic databases (AT, CY, DE, FI, LT, NL, TR)
- Feature structure of register/databases of geographical names (CZ, SI, HU, NO).

Common feature groups to all countries are settlements, administrative units, hidrography, orography/relief. Feature groups in some countries are also traffic and areas.

Geographical names attributes:

- 11 countries (AT, CY, FI, DE, HU, LV, NO, SK, ES, NL, TR) use feature coordinates
- 6 (CY, CZ, FI, FR, SI, NL) use name placement coordinates
- All countries except Lithuania have attribute of feature category
- 11 (AT, CY, CZ, FI, DE, HU, LV, NO, SK, SI, TR) have feature object ID
- 7 (FI, FR, LV, NO, SI, ES, TR) have map scale indicators
- 8 (CY, FI, DE, LV, NO, SK, ES, TR) have statistical classification
- 7 (CZ, HU, LV, NO, SK, ES, NL) have name status
- 4 (FI, NO, SI, ES) have language
- 1 (LV) has pronunciation
- 7 (AT, CY, CZ, HU, LV, SK, TR) have height
- 4 (CY, HU, LV, TR) have number of inhabitants
- 3 (FI, LV, SI) have map sheet number
- 1 (FI) has language status
- 2 (CZ, FI) have size + style characters
- 1 (DE) has gender
- 3 (HU, LV, NO) have name source and
- 3 (CZ, HU, LV) has variant names
- 2 (CZ, NO) has cadastral ID .

Formats:

- MS Word – 5 (AT, CY, FI, HU, LV)
- MS Excel – 11 (AT, CY, CZ, DE, FI, HU, LV, NO, SK, SI, ES)
- MS Access – 5 (AT, DE, LV, ES, TR)
- ASCII – 8 (CY, CZ, FI, FR, DE, SI, NL, TR)
- HTML – 2 (FR, LT)
- DXF – 2 (CY, LT)
- ESRI – 9 (CY, CZ, FI, DE, LT, LV, SK, SI, TR)
 - ArcInfo Coverage – 5 (CY, FI, DE, SK, TR)
 - ArcInfo Export – 5 (CY, FI, DE, SI, TR)
 - ArcInfo Shape – 5 (CY, FI, DE, SK, TR)
 - ArcInfo GeoDB – 1 (TR)
 - ArcGIS – 1 (CZ)
- Intergraph – 2 (AT, CZ)
- MapInfo – 3 (CY, FI, DE)
- Oracle – 4 (AT, CY, CZ, FR,)
- Microstation 95 – 1 (CZ)
- Visual Dbase – 1 (ES)
- Norwegian exchange format SOSI – 1 (NO)

Lists of exonyms:

- Available for 13 countries.
- Lists of exonyms are not available in Turkey. Latvia has materials for compiling such list.

Metadata:

- Available for 11 countries (AT, CZ, FI, HU, LT, LV, DE, NO, SK, SI, ES)
Six in ISO standard, other metadata standard are unknown.

Metadata of CY is in Greek, included in the Land Information System (not according ISO).

Germany has also in CEN standard, Slovenia in Dublin and Norway in OGC catalogue Services.

Access through Internet:

Geographical names data are directly accessed through Internet for 6 countries (DE, FI, FR, HU, NO, ES) and some of GN in another 2 countries (NL, TR).

Webserver search options:

- Complete names – **7** (FI, FR, DE, HU, NO, ES, TR)
- Letter combinations – **5** (FI, FR, DE, NO, ES)
- Coordinates/bounding box – **2** (DE, ES)
- Name categories – **3** (DE, HU, ES)
- Combinations – **1** (ES)
- County – **1** (HU)



Webserver standards for geographical names data:

- OGC WMS – 2 (FI, DE,)
- OGC WFS – 1 (DE)
- XML/SOAP – 3 (FI, DE, NO)
- HTML – 2 (FR, HU)
- GAZ - 1 (ES)

Test data:

Austria sent list of geographical names with attributes in Access file.

Cyprus has provided GN test data for three different products: ERM (e00), digital cadastral database (e00) and gazetteer (jpeg).

Sample of data for **Czech Republic** is as DBF, XLS, SHP and DGN format.

Finland has provided two alternative MS Excel tables including place and place name. To both alternatives are applied another MS Excel tables: feature types, language majority status, language official status, languages, municipalities, provinces and regions.

France has not provided test data.

Germany test data are in MS Access file and contain data in 17 tables of geographical names data model.

Test data:

Test data of **Hungarian** gazetteer are provided as MS Access file which additionally contains also county, sources and names type tables.

Latvian test data are provided in MS Access file – data are in 45 tables of database and contains all attributes. Descriptions of tables and attributes are in two other documents: SQL data types and toponymic database.

Lithuania sent data of topographic databases (e00) in which geographical names are one of the topographic feature attributes.

Lithuania also provided MS Access files with data for all tables and attributes of the data model of topographic base.

Netherlands sent data of topographic databases in which geographical names are one of the topographic feature attributes.

Norway has provided GN data in OUT and SQL files.

Test data:

Slovakian test data are provided as MS Access file. The table description and names of geographical features are enclosed.

Test data of **Slovenia** are provided as DBF, ASCII (txt file) and ArcInfo Export (e00) format.

Spanish test data are provided as MS Access file.

Test data of **Turkey** are provided as MS Access file. Description of feature types is enclosed in another table.