

## Common synopsis document of all Reference Group members

The synopsis document is a result of:

- Summarizing of detailed questionnaires and SI-EGN Final Report
- Checking reference documents
- Checking test data of geographical names.

First of all, reference documents provided by 15 countries are very different – from detail specifications and descriptions to more or less general information on geographical names. This is the reason that for some countries (SK and FR – because of language) synopsis base on detailed questionnaires filled in the beginning of 2005.

Test data are provided for 14 countries (except FR).

### Database

Not even one country does not have just analogue repository.

One digital stand-alone database have 9 countries (AT, CZ, DE, FI, HU, LV, SI, NL, NO).

Distributed databases are valid in 1 country (FR) and integrated databases for different scales in 3 countries (CY, DE, SK).

Mix between one digital stand-alone database and distributed database has another 2 countries (ES, TR).

Mix between one digital stand-alone database and integrated databases for different scales is the situation in 3 countries (FI, HU, SK).

In LT there are several (textual/tml) databases in the country not linked with geographic data and several geographic databases containing names as attribute values. Some of these databases can be linked to each other.

Two scenarios of names databases has just Germany.

Thus, prevailing models are one digital stand-alone database and integrated databases for different scales.

Data of geographical names of 11 countries (CZ, CY, DE, FI, FR, LT, NL, NO, SK, ES, TR) are linked to topographic, cartographic or cadastral data. Latvia has only some prerequisites for linking the toponymic data to the cartographic data.

In some countries (NL, LT) names are part – an attribute of topographic databases.

Operating system in use for software and hardware packages is in 13 countries MS Windows.

Other two (FR and SK) have LINUX.

UNIX is used together with MS Windows in 2 countries (CY, FI).

The most frequently used database package for storing the data is Oracle (9 countries - AT, CY, CZ, FI, DE, NL, NO, SI, ES). Follow MS Access (4 - DE, HU, ES, TR), SQL (3 - DE, FR, LV) and DBF (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).

All countries use different language codepages of ISO 8859 standard. MS Windows Codepages (1250 – 1257) are also used in 9 countries (CY, CZ, DE, HU, LV, LT, SI, ES, TR) and MS-DOS Codepage in 6 (CY, CZ, DE, HU, LV, TR). Unicode is in use just in 4 countries (DE, LV, ES, SK).

The 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.

In the cases that 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.

Different entity types (often geographical object, geographical name, map name) that are ID connected define data model of Finland, Norway and Slovenia. Each of entity type binds some attribute information of geographical names.

In the other side, Germany and Hungary data model include main table of the database contains different codes that are defined and characterized in different linked tables.

Geographical names are most often derived from the maps of middle scales: 1 : 20 000 (FI), 1 : 25 000 (FR, SK, SI, ES, NL, TR-populated places) and 1 : 50 000 (AT, HU, LV, LT, NO).

Names derived from large scale maps have 7 countries (1 : 5000 –CY, NO, SI; 1 : 10 000 – CZ, LT, SK, NL).

Four countries (DE, SK, SI, TR-other than populated) have geographical names for scale 1 : 250 000.

Some countries have geographical names derived from different scales, but usually one scale is preferred in reference documents.

#### Feature categorisation

Just 3 countries (SK, NL, TR) follow NATO feature categorisation standard. Majority (12 countries) have national feature coding structure.

National feature coding structures are divided into two types as:

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

In both names are divided into feature groups and further into several feature types.

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.

About formats in which geographical names are available there are the most differences between detailed questionnaires and reference documents and provided test data.

Countries have to check if statements in table below are correct:

- 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 are available for 13 countries.

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

Metadata is available for 11 countries (AT, CZ, FI, HU, LT, LV, DE, NO, SK, SI, ES). Six of them have metadata in ISO standard, for other the metadata standard is unknown.

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

Three countries have beside in ISO standard metadata also in one another standard: Germany in CEN, Slovenia in Dublin and Norway in OGC Catalogue Services.

Geographical names data are accessed through Internet for 6 countries (DE, FI, FR, HU, NO, ES).

Some of GN are available via Internet also in another 2 countries (NL, TR).

Geographical names data of Austria are not directly accessed via Internet, but it allows the search functionality in AMAP-online.

Webserver search options are:

- 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 are provided for 14 countries:

**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 of **Hungarian** gazetteer are provided as MS Access file which additionally contents 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.

**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.