

Minutes

of the

First Workshop of the EuroGeoNames Consortium and the Reference Group on Monday 23 October 2006, 13 – 19 hours and Tuesday 24 October 2006, 9 – 15 hours

at the

Bundesamt für Kartographie und Geodäsie, Richard-Strauss-Allee 11,
Frankfurt am Main, Germany

1 Executive Summary

The main points to note from the workshop are:

- The EGN Project Consortium elaborated the state of the art regarding the storage and maintenance of geographical names data, and presented the specific situation for the 15 countries referring to a ‘country-by-country document’ and a ‘synopsis document’. Both documents were accompanied by PowerPoint-Presentations. The Reference Group members were provided with handouts beforehand, summarizing the information for their country only. The documents/presentations will be uploaded to the EGN website (www.eurogeonames.com).
- The RefGroup members were requested to update the information concerning their database and to provide the EGN Project Consortium with amendments / corrections not later than **10 November 2006**.
- To gather new /additional information about the database tables, their structure, titles and content, the RefGroup members were requested to provide the Project Coordinator with test data not later than **10 November 2006**.
- Within the “Breakout sessions” the RefGroup expressed general ideas and proposals for mandatory and optional fields for a harmonized EGN data model. It will be used as a starting point for its development. The first draft EGN data model will be discussed within the 2nd Workshop with the EuroGeoNames Consortium and the Reference Group, which will be held very likely on **6/7 March 2006**.
- The concept for EuroGeoNames, especially regarding the services’ architecture and data flow was presented. The differences between an infrastructure of distributed and localized Web Feature Services (WFS) and one (huge) centralized WFS service connecting all distributed databases were explained, security issues, advantages and disadvantages, and maintenance issues were discussed. A proposal for an EGN services architecture and data flow will be presented within the 2nd Workshop with the EuroGeoNames Consortium and the Reference Group.
- With regard to variant names (‘exonyms’) it was stated that a work package is specifically designed for tackling this issue. It was pointed out that the concept of the variant names’ (‘exonyms’) database will be developed in consistency to the general EGN data model.

2 Participants

No	Country / Organisation	Surname	First name
Reference Group			
1	Cyprus	HADJIRAFTIS	ANDREAS
2	Czech Republic	ŠVEHLOVÁ	IRENA
3		HNÍZDILOVÁ	LENKA
4	Spain	ALCÁZAR	ADELA
5		AZCÁRATE	MARGARITA
6		ALONSO	JOSÉ ÁNGEL
7	Finland	LESKINEN	TEEMU
8		MYLLYMÄKI	TARJA
9	France	CALVARIN	ELISABETH
10		MAJKA	JEAN-SEBASTIEN
11	Hungary	POKOLY	BÉLA
12		MIKESY	GÁBOR
13	Lithuania	MARDOSIENE	DANUTE JANE
14		BECONYTE	GIEDRE
15	Latvia	GITENDORFS	ULDIS
16		STRAUTNIECE	VITA
17	Netherlands	BAKKER	NICO
18		STORM	MAARTEN
19	Norway	ANDERSEN	JOHNNY
20	Slovakia	SOKÁČOVÁ	PATRICIA
21		MÁJOVSKÁ	HEDVIGA
22	Turkey	TORUN	ABDULVAHİT
Consortium			
23	BEV, Austria	WURZER	WALTER
24	EDINA, United Kingdom	CORBETT	ANDREW
25	EuroGeographics, France	LOVELL	DAVE
26		JUST	LISE
27	ESRI, Germany	BUZIEK	GERD
28		SCHÄFER	AXEL
29	Utrecht University, The Netherlands	ORMELING	FERJAN
30	Geodan, NL	BLANKERT	ANNE
31	GeoTask, DE	SCHNEIDER	BERNHARD
32		PASSLACK	BERTHOLD
33	SMA, Slovenia	BRNOT	MARIJA
34	BKG, Germany	SIEVERS	JÖRN
35		ZACCHEDDU	PIER-GIORGIO
36		WALTER	HENNING
37		FALKE	HEDWIG
38		SPRADAU	JÖRGEN
Other Participants			
39	BKG, Germany	GRÜNREICH	DIETMAR
40	BKG, Germany	BENNAT	HEINZ
41	BKG, Germany	MAU	RONALD
42	BKG, Germany	ILLERT	ANDREAS
43	BKG, Germany	MORDHORST	RONALD

3 Attachments

All documents and presentations discussed within the workshop will be uploaded to the EGN website (www.eurogeonames.com).

4 Agenda

TOP	Topic title
1	Opening
2	Presentation of the concept of the EuroGeoNames project
3	Presentations of aims and expectations from the geoinformation market towards the EuroGeoNames project
4	General questions
5	Presentation of the synopsis of the state of the art of geographical names data information
6	Identifying the key technical issues (broad topics for discussion)
7	Exchange of ideas concerning User/business requirements
8	Approaching a consensus
9	Breakout session / Presenting the outcomes
10	Final discussion
11	Next steps & closing

5 Minutes and results of the discussion

Minutes	Action
<p>1 Opening</p> <p><i>Sievers</i>, as EGN Project Co-ordinator, opened the workshop and emphasized the objectives of the workshop.</p> <p><i>Grünreich</i>, President of the Federal Agency for Cartography and Geodesy (BKG), welcomed the workshop participants at the premises of BKG. His welcoming address was focussed on emphasizing the importance of geographical names in the context of national and international spatial data infrastructures (SDIs). He explained the measures which have been already taken in Germany concerning the establishment of the spatial data infrastructure GDI-DE® and he pointed out the importance of the EuroGeoNames project in the context of setting up a European SDI.</p> <p>The PPT-presentation will be uploaded to the EGN website (www.eurogeonames.com).</p>	BKG

<p>2 Presentation of the concept of the EuroGeoNames project</p> <p><i>Zaccheddu</i>, as EGN Project Manager, gave an overview of the concept developed for the project EuroGeoNames and reported about the current status of the negotiation process with the European Commission. In a nutshell, he stressed that EGN is aiming at developing a geographical names infrastructure and services by aggregating the national sources of geographical names data supporting their national languages. He pointed out how the Reference Group will be involved in the development and the review processes, displayed in the spiral engineering process.</p> <p>The PPT-presentation will be uploaded to the EGN website (www.eurogeonames.com).</p>	BKG
<p>3 Presentations of aims and expectations from the geoinformation market towards the EuroGeoNames project</p> <p><i>Buziek</i> presented general aims and expectations from the geoinformation market towards spatial data. He pointed out how the EuroGeoNames project could play an important role in populating the market with valuable spatial data, recognizing that geographical names are one of the most important spatial data components.</p> <p>The PPT-presentation will be uploaded to the EGN website (www.eurogeonames.com).</p>	BKG
<p>4 General questions</p> <p>None</p>	
<p>5 Presentation of the synopsis of the state of the art of geographical names data information</p> <p><i>Brnot / Wurzer</i> had worked out and elaborated the state of the art regarding the storage and maintenance of geographical names data for the Reference Group countries. <i>They</i> presented the specific situation for the 15 countries by referring to a 'country-by-country document' and a 'synopsis document'. Both documents were accompanied by PowerPoint-Presentations. For all countries information was compiled about:</p> <ul style="list-style-type: none"> - the current database structure, - the sources used to populate the databases, - the feature categorization and the attribute information to the geographical names, - the data exchange formats, - the availability of metadata, 	

<ul style="list-style-type: none"> - the possibility to have access to the data through the internet, - the existence of a list of exonyms and - the provision status of test data files. <p>The Reference Group members were provided with handouts before the workshop started, summarizing the information for their country only. The documents/presentations will be uploaded to the EGN website (www.eurogeonames.com).</p> <p>The RefGroup members were requested to update the documents concerning their databases and to provide the EGN Project Consortium with amendments / corrections not later than 10 November 2006.</p> <p>Within this context it was clearly stated that the EGN Consortium needs not only describing information. In order to gather new/additional information about the database tables, their structure, titles and content, the RefGroup members were requested to provide the Project Coordinator with test data files not later than 10 November 2006.</p> <p>Within the discussion, language issues, e.g. ISO codepage conversion to Unicode, or metadata standards like ISO 19115 were touched on. Furthermore it was clarified whether or not the geographical names databases are linked to topographic data and whether the attribute information is related to point features only or to vector information (point, lines, areas).</p>	<p>RefGroup</p> <p>RefGroup</p>
<p>6 Identifying the key technical issues (broad topics for discussion)</p> <p><i>Corbett</i> presented issues concerning the envisaged EGN data model. He pointed out that on the one hand the model could be a more sophisticated one (of which not all fields have to be filled) that on the other hand, however, a minimum number of denominators (e.g. a unique object ID, the geographical name, the feature type and the location information) have to be mandatory. Within this context he explained the differences between a services' architecture and data flow of distributed and localized Web Feature Services (WFS) and one (huge) centralized WFS service connecting all distributed databases. Advantages and disadvantages as well as maintenance issues were discussed. He pointed out that the provision with test data files would be crucial in order to start with the development of a harmonised EGN data model and to determine the map entities and the minimum WFS capabilities (query by object name and/or feature type and/or location).</p> <p><i>Mordhorst</i> gave a presentation about security issues to be considered in a distributed services' architecture. Questions like "Distinguish between different levels?" or "How can the countries secure their data?", "Do we accept an anonymous user? <-> What about an expert user?" were raised. He emphasized that WFS services have to be language-independent and that there is almost always a way for a solution. Hence, he stimulated to start with discussions regarding this matter as soon as possible, because "security issues" would be a general problem for all countries providing data via the Internet. As a</p>	

current state-of-the-art/starting point, he mentioned that different classes of users could be defined, and that e.g. a 'https channel' -> for normal user and 'Virtual Private Networks' -> for commercial applications could be considered.

The following broad topics were discussed more intensively:

Data Access and services' architecture:

It was clarified that a WFS is a piece of software, to be installed for running on the countries' system. Technically, so-called SQL-queries would be send against the countries' databases. With regard to the suggestion to harvest the content of the countries' databases for improving the performance of the EGN infrastructure, it was mentioned by the Project Consortium that, due to political/strategic constraints, a distributed services' architecture is mandatory for EGN. It was agreed upon that EGN should aim at WFSs to be adapted to the countries' databases. Therefore, country-specific circumstances have to be taken into account (e.g. written API in Norway, which connects the "shadow database"; or in Spain the already implemented WFS which would be connected to EGN, instead of implementing an additional EGN service at the client's database).

As for the GI market in general, issues of data quality were considered to be very important. Here metadata would be very helpful and crucial.

With regard to metadata it was generally stated that the development of a metadata profile for geographical names (according to the ISO 19115 standard) will be part of a EGN work package (WP3).

Feature catalogues:

Issues were discussed related to objects named in several languages (e.g. in Finland), to unique object identifier as well as to schema translation and conversion from the national data models to the EuroGeoNames data model.

As a general remark it was suggested to consider and recall the EuroGeographics (EGG) products EuroRegionalMap (ERM), EuroGlobalMap (EGM) and EuroBoundaryMap (EBM, formerly called SABE) while developing the content and level of detail for the EGN data model.

It was clarified that currently an overall model of EGG products is not yet available, but that the respective EGG products could provide a draft feature catalogue (but, which is not yet adopted) for EGN. It was explained that the framework is available in English, but that definitions have still to be agreed upon. It was suggested to try to find the feature types first in EGG products (e.g. in ERM), then to investigate the Nato FACC specifications before creating new feature types. In the ERM feature catalogue, landscape/orography features are included as a text string in "Named location", but generally no feature types/classes for landscape/orography features have been defined yet. Some countries (ES, NO, FR, NL, DE, etc) are already dealing with landscapes/orography. Hence, in this case EuroGeoNames could give valuable input to EGG products.

With regard to questions related to a 'dictionary on generic terms' it was

<p>stated that a 'drop-down' menu in the local language for objects and symbols for features could be facilitated within the envisaged application?</p> <p><u>Variant names ('exonyms')</u>:</p> <p>With regard to variant names ('exonyms') it was stated that a work package (WP5) is specifically designed for tackling this issue. An online-service for editing/revising variant names by the European experts themselves would be designed within this work package. It was pointed out that the concept of the variant names' database will be developed in consistency to the general EGN data model.</p>	
<p>7 Exchange of ideas concerning user/business requirements</p> <p><i>Ormeling</i> presented some assumptions as well as general ideas regarding user and business requirements. He opened the discussion with the RefGroup asking for amendments, general remarks and further suggestions.</p> <p>He announced the workshop on user/business requirements, which will be held at the University of Utrecht, on 15 January 2007. The RefGroup would be kindly invited to participate in the workshop (unfortunately without any reimbursement for travel and subsistence).</p> <p>He explained that the outcome of the workshop would be a final report on user/business requirements, which will be a major basis for the further development of the EGN data model as well as of the services' architecture. He stated that as a starting point, potential applications for specific users or firms would be analysed (e.g. routing planner).</p> <p>Within the discussion, it was pointed out that all assumptions stated by <i>Ormeling</i> should be displayed with a question mark, because, according to the 'Declaration of participation' which had been signed by all RefGroup members, not "<i>everything is available, accessible and possible</i>", thus the RefGroup agreed upon to provide data for single inquiries only.</p> <p>ESRI stressed that the EGN mission/goal (within the 30-month project duration) has to be assessed against the EGN vision. This would be very important for the GI market in order to define a clear product management (communication process) for the customer/user to bring the 'EGN product' in the right shape. It means to define the customer's data availability and accessibility.</p> <p>Street names or historical names should be considered in the EGN vision, but not in the first project approach.</p> <p>The outcome of the workshop on user/business requirements will be presented at the second workshop of the Consortium with the RefGroup, very likely to be held on 6/7 March 2007.</p>	RefGroup
<p>8 Approaching a consensus, resolving the issues (reprise from previous day)</p> <p><i>Sievers</i> started with presenting a resumé of the first day:</p> <p>(1) Data access: EGN <u>and</u> countries nationally to decide on security issues /</p>	

<p>WFS infrastructure</p> <p>(2) Architecture: Trend to distributed WFS implemented at location</p> <p>(3) Feature catalogue: Trend to ERM, Nato FACC, feature catalogue and dictionary</p> <p>(4) Variants ('exonyms'): feature catalogue and dictionary (see (3)), online-editing-service</p>	
<p>9 Breakout sessions and outcome from the breakout session</p> <p>For the breakout sessions four groups were established in order to discuss more intensively:</p> <ol style="list-style-type: none"> 1) minimum requirements for the EGN data model and the EGN output for the EC-funded project duration of 30 month 2) the EGN vision <p>Outcome from the breakout session:</p> <p><u>Group 1:</u> EGN has to build a harmonized model, but a very heterogeneous situation in all countries will be faced.</p> <ul style="list-style-type: none"> - Name, point information (location), feature type <p>Not allow queries which cross boundaries. Very simple feature typing. Feature catalogue: use the FACC Geodetic reference frame: use ETRF89</p> <p><u>Group 2:</u> General issues:</p> <ul style="list-style-type: none"> - what should be the information to be provided by EGN? - Starting point: small scale, 1 : 250.000 - ISO 19112 - geographic identifier - direct and indirect referencing (e.g. coordinates and administrative units) <p>Feature classes:</p> <ul style="list-style-type: none"> - <u>Administrative units</u> feature classes according to EuroBoundaryMap -> NUTS codes - <u>Hydrographic features</u> feature classes according to EuroRegionalMap, FACC ??? - <u>Transportation</u> feature classes according to EuroRegionalMap, FACC, EuroRoads ??? - <u>Populated places</u> feature classes according to EuroRegionalMap, FACC ??? - <u>Orography</u> feature classes according to ES, DE, FR, LV, NO, etc. - <u>Vegetation & soil</u> feature classes according to EuroRegionalMap, FACC ??? - <u>Points of interests (Miscellaneous)</u> 	

<p>feature classes according to EuroRegionalMap, FACC ???</p> <p>1) Minimum requirements: Gazetteer: Name, Feature type, location <i>Mandatory Object/Feature</i> (-> only geographical attributes yet):</p> <ul style="list-style-type: none"> - Unique object ID - Name - Feature type - Coordinates - Administrative units <p>2) Vision of EGN data model: Issues for consideration, differentiation of attributes to the name:</p> <ul style="list-style-type: none"> - <i>Geographical attributes</i> (unique object ID, feature type, spatial reference, object history/status) - <i>Linguistic attributes</i> (languages -> ISO codes, pronunciation -> IPA codes or audio-files, gender, short form, genitive form, etc.) - <i>Class name</i> (main or preferred name, variant name, historical or former name, etc.) - <i>Etymology</i> (origin) - <i>Status</i> (official/administrative (FR), standardized, not standardized, not available, ...) - <i>Source</i> - <i>Other attributes</i> (erroneous, associated adjective,...) <p><u>Group 3</u></p> <p>Not to use simple hierarchical structures for modelling, but diagrams that can express complex relationships</p> <p>Conceptual model – two main parts:</p> <ol style="list-style-type: none"> 1. real world object 2. name <p>Relationship between real world object and name is 1 to n</p> <p>Attributes to real world object:</p> <ul style="list-style-type: none"> - Location (ISO19111: by coordinates or ISO19112: by geographic identifiers, both might refer to secondary data sets) - Feature type - Unique ID <p>Attributes to name:</p> <ul style="list-style-type: none"> - name - language - metadata - status – attribute values from UNGEGN list? - importance of the feature/name -> Map representation, - normalization – romanization? 	
--	--

How do we treat different alphabets? Entries in different character sets? -> Unicode
 Cross border features, e.g. Danube, Alps, ... -> only one real world object!
 Feature catalogues – hierarchy of feature catalogue, e.g. river might be part of hydrography and transportation - > multiple relations
 EGN feature catalogue <-> national feature catalogues -> translation of feature types!

Group 4:

Conceptual question: main entry is the name? Or the feature?

No decision about what will be the preferable way.

Classes attached to a gazetteer:

- object name (language, name)
- gender
- script
- narrative description
- pronunciation

Status and time period not mandatory. Addresses should be optional.

Shouldn't the addresses be part of the object name? Variant of indirect identifier?

Location -> Geometry type: point, line, area, bounding box, etc. ;

Geometry: coordinates, bounding boxes, etc.

Have the option to attach multiple locations.

There was no extensive discussion about metadata, but it was considered to be very important and, therefore, mandatory. Information is needed about quality.

As for object types already existing standards should be followed.

10 Final discussion

Schneider introduced the final discussion by highlighting that the Project Consortium would work together with the Reference Group on the EGN data model definition.

He pointed out main issues, opportunities and current barriers of the envisaged EGN project by illustrating query examples raised for serving different purposes .

He stressed that commercial companies do expect homogenous quality data, complete pan-European coverage and high availability.

Within the final discussion mainly the following achievements were identified:

- It was seen to be the major achievement of the workshop to have gained a better common understanding of which geographical names data is available in the countries. By updating the documentations and providing the Project Consortium with test data files, the EGN data model definition would be supported satisfactorily, and possible functionalities / limits for

<p>potential use cases would be recognized in good time;</p> <ul style="list-style-type: none"> - It was recommended to wait for the outcome from the workshop/work package on user/business requirements to give answers about possible use cases/applications and about what the final products of EGN would be; - It was recommended to prepare toponymic guidelines by each country in accordance to resolutions of the United Nations Conferences on the Standardization of Geographical Names. The general framework for toponymic guidelines can be downloaded from the UNGEGN website (http://unstats.un.org/unsd/geoinfo/about_us.htm); - It was stressed that the management of multi-lingualism as well as variant names would be part of the EGN data model like the ‘status’ of the name (official, in-official, etc.). Some countries have planned to share experiences regarding this matter in due course (Latvian with Finish and Spanish colleagues). <p><i>Sievers</i> emphasized that the input given by <i>A. Illert</i> to the discussion about how the issues of common feature catalogues/typology will be addressed in future was highly appreciated and will be continued. This input is very important considering that he is the chair of the EuroGeographics WG on “Data specifications” as well as a member of the Inspire Drafting Team on ‘Data specifications’. By that interdependencies between EuroGeographics, its members, EuroGeographics’ EuroSpec programme and the EGN project could be recognized and synergies with the Inspire process be assured.</p>	
<p>11 Next steps and closing</p> <p><i>Sievers</i> presented the five planned upcoming workshops (comprising the second Workshop, which will be conducted explicitly with the Reference Group) as well as of the contributions requested from the Reference Group members/institutions.</p> <ul style="list-style-type: none"> - WS2 User/business requirements (organizer Ormeling), January 2007 - WS3 with Reference Group (organizer BKG), March 2007 (costs for travel & subsistence for RefGroup are foreseen) - WS4 Business model (organizer GeoTask), June 2007 - WS5 Implementing Pan-European EGN Web services (organizer EGHO) (likely) November 2007 (costs for travel & subsistence for RefGroup are foreseen) - WS6 Closing (organizer EGHO), December 2008 (costs for travel & subsistence for RefGroup are foreseen) <p><i>Sievers</i> finally thanked all participants for the fruitful discussion, for the good working atmosphere and for the very positive attitude towards the EuroGeoNames project.</p>	