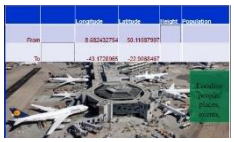
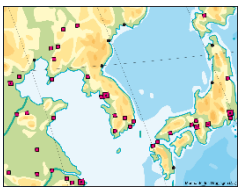
The role of UNGEGN

I thank GGIM-Europe for the opportunity to explain to you the role of the other subsidiary body in the field of geospatial information management that operates, next to you, under the aegis of the United Nations Statistics Department

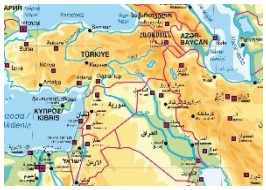
To do so, I start with your journey to Brussels – imagine that at the airport the



flight departure sign boards would use geographical coordinates instead of geographical names – only the geodesists amongst you might have a *clue* where the flight shown here starts and where it intends to arrive. Even when the third dimension and some demographic data would be added to the board, the location of the departure and arrival points will remain opaque (opeek) to most of you. It will remain an enigma until I add their geographical names. Names are central in providing access to geospatial information, be it for communication, for real estate, for news platforms or for emergency operations.



Using more points won't make any difference. Seeing this map with city locations, you must agree with me that geographical names are an essential element of geospatial data. There would be alternatives to refer to the cities portrayed here, such as, again, geographical coordinates, but those would dehumanize. Geographical names are the primary tool to distinguish geospatial objects one from another, but even if they are available, they not always are rendered in a form everyone would understand. So, we would need *conversion systems* in order to render these names in a form more accessible to us. Rendering them in our Roman alphabet would call for a romanization system, rendering them in Arabic, Cyrillic, Devanagari or Mandarin would call for other conversion systems. Even then, that would not be enough, as the pronunciation of the letters in the Roman or Cyrillic or Arabic alphabet is different in different language areas or countries. So, we would need information on the *pronunciation* as well: first on the system to use for recording the pronunciation, consecutively for rendering this pronunciation in a way suitable for the target audience. This all sounds like a task once done, should be sufficient for eternity, but that cannot



be the case. Let us look at the Near East: as was the case for East Asia this map is mostly inaccessible when the names are rendered in the local official scripts. So, we have to make this map accessible by converting the local names in their

local scripts (Arab, Greek, Hebrew, Armenian, Georgian and Cyrillic) to the Roman alphabet. But if we do so, we are confronted with the existence of *conventional names*, that is customary names with which geospatial objects are designated by the international community. The names on view here may look familiar to you, but they are *not* the official names. Many of them go back to Antiquity, we have kept these old name forms in Europe, even when in the Near East other civilizations came up and changed them. So, it will be no wonder to you that the official versions of the current names, rendered

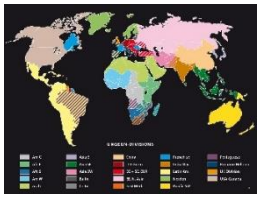
into the Roman alphabet following official conversion systems, will look totally different. Standardizing names and standardizing conversion systems is not enough – one also has to make sure that the standardised names files are being kept up to date.

So, with this series of images I try to show what UNGEGN is for: we work out methodologies or best practices for the standardization of geographical names. Our aim is *univocity*, that is the existence of only one standardized official name version for every topographic object in each writing system. Apart from names standardization, we go in for the standardization of *conversion systems*, ideally for systems that are reciprocal, so that when one changes a name from Arabic into Roman alphabet and back again, the result would be identical to the original spelling. We work out *best practices* how to store the standardized names and making them available in gazetteers, names servers, or any geospatial data set using standardized names within a spatial data infrastructure that can be consulted world-wide. And we keep stressing the need for frequent updates of these toponymical databases in order to avoid the discrepancies we have seen for the Near East between outdated names and official names.

In UNGEGN we don't have the mandate to decide on individual names: we cannot prescribe which names should be used by the international community for specific objects, let alone which names should be used nationally. Each nation is sovereign in deciding what names to use. We establish procedures following which countries may settle their toponymical claims.

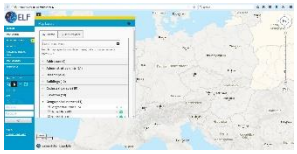
Just to give you an idea of our structure: individual countries bond together in *divisions* on the basis of their proximity or their linguistic kinship, in order to tackle joint problems or issues or

projects. If they have experts with specific expertise in certain fields, these would cooperate in UNGEGN *working groups*. We have topical Working Groups on: country names, on data files, terminology, romanization systems, exonyms, pronunciation and on geographical names as cultural heritage. More organisational are the working groups on Publicity and on the Implementation of the resolutions. Our Group of Experts is facilitated by a bureau and a secretariat, and we have a special task team for Africa, as well as a coordinator for *toponymical guidelines*, that is a publication to be issued by every country for the benefit of foreign editors, showing the intricacies of the use of that country's geographical names. Right now, we work on a five-yearly basis, that is with a conference every five years, but we are now considering to change that to a four-yearly basis in order to be more compatible with other UN-organizations.



The UNGEGN divisions do not match those of GGIM, as linguistics plays a part in them. Because of the overseas distribution of English, French, Spanish, Portuguese and Dutch, the distribution of the divisions gives a chequered image, also because each country can be a member of more than one division.

To give you an idea of the activities of divisions, I take the Dutch- and German-speaking division, of which Belgium, Austria, the Netherlands, Switzerland, Suriname and South Africa also are members. This Dutch and German-Speaking Division has been instrumental in realizing the *EuroGeoNames* project for the European Union and for EuroGeographics, and is now providing the geographical names for the European Locator Framework (ELF) Geolocator service. It also is regularly organizing seminars on names issues in Western Europe, and has specialized in organizing toponymy courses worldwide. A group of lecturers from the Dutch- and German-Speaking division is just back from a course in Brazil.



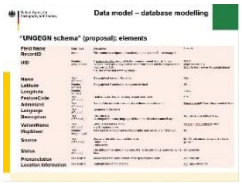
In our 50-year existence we have developed glossaries and technical manuals for national standardization of geographical names, that can be downloaded from the UNGEGN website. The most important part of our website probably is the section developed by our *Working Group on Romanization systems* which has the UN-approved conversion systems for most of the non-roman scripts and writing systems like Arabic, Cyrillic, Chinese, Greek and Hebrew. This is the site most consulted by toponymists all over the world.

A multilingual, multiscriptual geo-referenced geographical names database is being developed by UNGEGN, with input from the UN Cartographic Section and the UN Geographic Information Working Group (UNGIWG). Through the web, database users can access short and full names of countries, their capitals, and the major cities. Authoritative city endonyms are provided mainly by national name authorities and *sound files* are being added to assist users with pronunciation. As a useful reference tool for geo-information management, this *UNGEGN World Geographical Names Database* will continue to be developed and improved, and updated on an ongoing basis.

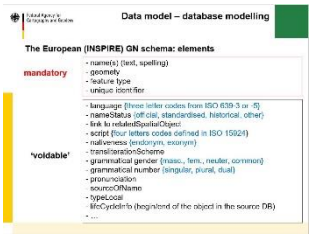
The link between geographical names and spatial data infrastructures was recognised early amongst us, and was formally acknowledged by ECOSOC with the acceptance in 2002 of UNGEGN Resolution VIII-6, proposed by Germany, on the Integration of geographical names into national and international spatial data infrastructures.

We are currently developing an advanced toponymy manual, which is a sequel to the web courses in toponymy, created in English, Spanish and French. There is only a small number of toponymists to be trained in every country, so that setting up university courses does not seem feasible in most countries, and therefore we also have an outreach programme, and toponymy courses are given all over the world. So, we sustain this *capacity building* activity about all name-related issues such as multilingualism, exonyms and endonyms, transliteration as well as all technical database or data infrastructure-related issues in order to create names sources accessible through national and international spatial data infrastructure (like INSPIRE in Europe). UNGEGN has also developed manuals teaching names standardization techniques, that go from names collection in the field to office processing of those names to building the databases.

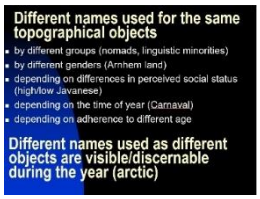
UNGEGN has provided substantial input to INSPIRE geographical names data specifications. The INSPIRE implementing rule with the INSPIRE geographical names data specifications reflect UNGEGN's concepts. In addition to providing a basis for the interoperability of spatial data in INSPIRE, the data specification development framework and the thematic data specifications can be reused in other environments at local, regional, national and global level contributing to improvements in the coherence and interoperability of data in spatial data infrastructures. The main value of the INSPIRE



geographical names model is a simple yet flexible structure that allows geographical names to be used as an attribute of a spatial object, either modelled within the geographical names theme or in any other theme of INSPIRE. The possibility of linking more names with the same named places gives the opportunity to integrate minority languages and exonyms, which are an important contribution to European multilingualism.



In Europe, topographers hardly leave their offices nowadays for names collection, and here we are experimenting for the update of our names with volunteered geographic information. But outside Europe fieldwork is still necessary for names collecting, and there one is confronted with different usages regarding geographical names; we find for instance



-Different names used by different groups (nomads, linguistic minorities) for the same objects

-Different names used by male and female inhabitants for the same objects (as in Arnhem land in Australia)

-Different names used depending on perceived differences in social status (high/low Javanese)

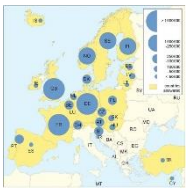
-Different names used depending on the time of year (Carnival)

-Different names used depending on the adherence to different age groups. Now that we are becoming dependent in Europe on volunteered geographical information for getting informed on new names or the disappearance of old ones, we have to take account of the possible age bias in this information.

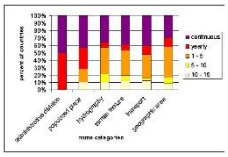
-Finally, different names may be used as different objects may be visible or discernible throughout the year, such as stretches of open water in the Arctic

The changing age structure but also the accelerating urbanization process make us aware of the danger of losing our *traditional rural names*: names for arable fields, for hills and brooks, or for microtopography, that, because they are not used any more in daily life, tend to be forgotten. This *cultural heritage*, that tells us so much of the original conditions of the land, of the vegetation and of the colonizers, should be preserved, and we are developing programmes to that effect.

Quality control is one of the most pressing aspects of toponymy and will become a major issue in UNGEGN in our work programme. We should be auditing our names databases, checking spellings and bringing them up to date if necessary. If we have different authorities in our countries in charge of different types of geographical names, these names should be matched, it should be checked whether the same name elements are spelled similarly. Statistical, administrative and topographical names should be compared and adjusted one



to another. When we look at **quality control** at a European level, we would like to have a similar density of names collected, processed to similar standards in our European names database – this slide is from a survey we did 10 years ago, but I do not expect the present image to be much different – so there is a clear discrepancy in the names density in our current national databases. The number of names for Cyprus seems to be similar to the number of names for Turkey in their national names databases. Some countries in Europe do not have separate name databases, but instead, the official names are included as attributes to the named places or objects in the digital landscape models of these countries.



When we look at the graph, about half the European countries only update their names databases every 5 years, except for the administrative data which are updated continuously. If we look at the institutions responsible for standardizing geographical names, there is a mix between those linked to the national mapping or cadastral administration, specially designated geographical names boards and statistical agencies. Some countries, like the Netherlands and Britain don't even have a national names board at all. Coloured bands refer here to divided authority, different institutions in charge of the standardization of different name categories. Of course, this need not be detrimental to quality, but at least it does not make quality control easier. This next map shows the source map used originally by the European countries for the national database or gazetteer, and it can be seen that for a country like Cyprus the names have been taken from a map at the scale 1:1000, while for others they have their names from a 1:100 000 map. So the names density and the updating practices of the names databases are rather unequal throughout Europe, while the authority to standardise the different names categories can be divided over several institutions, often under different ministries. This is contrary to what we see as the ideal situation, that is the

provision of an even, homogeneous layer of names, that are regularly checked. And here we hope very much to engage with GGIM-Europe. We hope, by having regular contacts with you, to raise the awareness and the sensibility amongst those in charge of the national topographic, cadastral and statistical agencies, regarding the necessary names standardisation processes. We should find ways of exchanging ideas and to gain a common understanding of the issues related to the – in GGIM-speak - *fundamental data theme geographical names,* in raising the quality level of the names databases. Auditing of these national databases is required, in order to bring them to the same standards. And that requires a closer cooperation between national names boards, national statistical offices and national mapping agencies. In closing, I would like to refer to the UNGEGN-UNGGIM relationship document titled ”*Building Bridges*” and the common principles stated in that document, that will be dealt with and hopefully accepted by both organizations in the separate conferences they will be having this August in New York city.

Thank you