GENERALISATION WITHIN NMA’S IN THE 21ST CENTURY

J.E. Stoter
ITC, International Institute for geo-information science and earth observation
P.O. Box 6, 7500 AA Enschede, the Netherlands
stoter@itc.nl

Abstract
With the automation of production processes within National Mapping Agencies (NMA’s), the question arose if generalisation can be performed automatically. Many research results have been achieved on generalisation, although there are many aspects that still need more attention, such as contextual generalisation, DTM generalisation and map names generalisation. Are research results implemented in practice and what developments are still needed to support automatic generalisation processes within NMA’s? These questions will be addressed in this paper based on the outcomes of the generalisation workshop, organised on 31 March and 1 April 2005, ITC, Enschede (the Netherlands). Generalisation processes as well as policies on generalisation within the twelve participating NMA’s will be described. From these experiences it can be concluded that all selected NMA’s are faced with the question how to implement automatic generalisation (or at least as automatic as possible) within a renewed production line. Some NMA’s have already made major steps in this process, while other NMA’s are still at the beginning.

1. INTRODUCTION
Traditionally National Mapping Agencies (NMA’s) are responsible for the production of topographic maps at different scales. From the beginning of the seventies map production within NMA’s was increasingly supported by maintaining the lines, points and polygons representing topographic objects, as well as symbols and map names in a digital environment. In some cases this was done by storing the symbology directly with the topographic information and in other cases this was done by the separation between the actual topographic information (Digital Landscape Model) and the way the information appears on a map (Digital Cartographic Model). Most NMA’s introduced the concept of generalising small scale maps from large scale maps already before the availability of digital topographic datasets. Updates were carried through at the largest scale and based on the updated information the cartographer propagated the updates into the smaller scales by applying generalisation rules in a sequential order. The definition of generalisation rules such as “if buildings touch or overlap roads, they should be replaced to a distance of at least 0.2mm from the road, unless the real distance is less than 6 meters” standardised the way cartographers generalised maps. This generalisation process was initially only partly automated: the large scale map was projected on a screen and the cartographer created maps at smaller scales by applying the generalisation rules interactively. With the availability of digital topographic information in a vector based database-environment, the question arose if information at smaller scales can be directly and automatically derived from large scale topographic datasets. In research much has been achieved on some main aspects of generalisation. Do NMA profit from these achievements and what research gaps still has to be filled to better support generalisation within NMA’s? To answer these questions a workshop was organised at ITC, Enschede, on 31 March and 1 April, 2005 with two main goals: a) to get an overview of the state-of-the-art on generalisation within NMA’s and b) to get insight into the topics that still need to be addressed by research to better support generalisation processes in NMA’s.

The workshop was attended by twelve NMA’s: Belgium, Catalonia (Spain), Denmark, France, Great Britain, The Netherlands, Ireland, Switzerland, Sweden, and three Bundesländer from Germany: North Rhine Westphalia, Baden-Württemberg and Lower Saxony. In this paper the results of the workshop are discussed. We have tried to make a clear comparison between the different NMA’s. However, because of different scales, different definitions of similar terms, different terms for similar meanings, different technological and organisational structures of the NMA’s, this was not easy to do. Therefore we start with a description of generalisation practice within the selected NMA’s (section 2). In section 3, we conclude on the experiences of the selected NMA’s by answering the two main questions from the workshop. For the state-of-the-art on generalisation, see also (EuroGeographics, 2005) and (Ruas, 2001). Note that this paper specifically focuses on datasets maintained by National Mapping and Cadastral Agencies (NMCA’s) for topographic mapping and not on other datasets such as cadastral data, orthophotos and DTMs.
2. GENERALISATION WITHIN NMA'S

2.1 Belgium

IGN Belgium (IGN, 2005) maintains vector datasets at scale 1:10k, 1:50k, 1:100k and 1:250k. The Top10v-GIS (1:10k vector database, maintained as ArcInfo coverages divided into map sheets) is currently created based on photogrammetric survey (stereoplot), field survey (identification on aerial photographs or, specific for the roads, on 1:10k analogue maps) and plans (for public works). It will be finished by the end of 2006. The first digital edition of Top50v-GIS (1:50k vector database) was produced between 1994 and 2001. Because of different production cycles, the 1:10k dataset could only be used to produce half of the 1:50k scale dataset. The other half was based on the 1:25k old analogue maps. The whole edition was loaded in a seamless ESRI geodatabase maintained in Informix. Since 2002 the update of the 1:50k is performed in the field, which means that 100% interactive generalisation takes place from real world objects to the 1:50k dataset during this scale specific data collection. The second digital edition of Top50v-GIS will be finished in 2007. The first edition of the 1:250k dataset was produced by digitising the old 1:50k analogue maps. It is now being updated separately every year.

The 1:100k (Top100v) is currently being derived partially automatically from 1:50k using ArcGIS, with self-developed algorithms in AML (ArcInfo Macro Language) while the buildings outside cities are generalised using the Laser-Scan Lamps2/Agent prototype. The production of 1:100k started in 2001 and will be finished in 2006. The generalisation process consists of five steps: 1) selection, assimilation and displacement of roads, hydrography and railway network, 2) generalisation of point features, 3) generalisation of land use, 4) generalisation of buildings, and 5) generalisation of administrative data. After each step the result is checked and, if necessary, corrected interactively.

In the coming years all the datasets will be available for the whole country. IGN Belgium is in the process of deciding on future production line. The fundamental question they are faced with is whether to maintain one reference database at scale 1:10k and generalise as automatically as possible the datasets at scales 1:50k and 1:100k or keep the different datasets while only generalising the updates (as automatic as possible). The integration of Top10v-GIS and TOP50vGIS is now being studied. One of the issues identified is the difference of update date, which complicates the feature linking process.

2.2 Catalonia, Spain

At this moment ICC, Catalonia (ICC, 2005) produces and maintains four topographic vector databases: 1:5k, 1:25k, 1:50k and 1:250k. Apart from these four vector databases, digital maps at scale 1:10k, 1:50k, 1:100k and 1:250k are maintained. Although datasets are maintained in map sheets, geometric, topologic and semantic continuity is warranted. Current generalisation workflow in ICC is illustrated in figure 1.

![Diagram](image-url)

Figure 1: Current generalisation workflow in ICC, Catalonia
The 1:5k database (completed in 1995 and updated between 1997 and 2005) contains 2.5D data and is a GIS oriented database. From the 1:5k database the 1:5k map is generated using automatic symbolisation. The 1:10k map is generalised from the 1:5k database by applying both automatic (CHANGE for buildings and ICC software) and interactive generalisation. The database 1:25k is obtained by applying generalisation (using CHANGE and ICC software and human interaction) to the 1:5k database, based on generalisation specifications. The 1:25k database will be updated separately from the 1:5k database. The 1:25k map is obtained by automatic symbolisation of the 1:25k database. The 1:50k database was defined and filled separately from the 1:5k database, although the data models are consistent. The first edition of the map at scale 1:50k (Mapa Comarcal) was derived from the 1:50k database in 1991. However, it is updated separately from and more frequently than the 1:50k database since it requires a lot of work to produce the map from the vector database. Since the up-to-date requirements for the maps are larger than for the databases, the emphasis is on updating the map first. This leads to the unusual situation that the cartographic output is newer than the underlying database. The map at scale 1:100k is currently being derived by generalisation from the 1:50k map using ICC software combined with human interactivity. The first edition of the 1:250k database was derived from the 1:50k database. Now it is updated separately from the 1:50k database. The first edition of the map at scale 1:250k was obtained from the 1:250k database. However, as for the 1:50k scale, it is updated separately from and more frequently than the 1:250k database.

In conclusion, current generalisation within ICC consists of automatic generalisation processes, with human interaction afterwards. However, more automatic generalisation processes would save a lot of human interaction. ICC is currently reconsidering its workflow since it is too expensive to update all the datasets while maintaining the coherence between them. The new workflow, in which the 1:50k database will be eliminated, will take the following requirements into account: 2.5D generalisation, map name generalisation, contextual generalisation and maintaining links between original and derived datasets. It should be noted that digital information (vector and raster), produced by ICC is free, only the 1:5k database has an inestimable price.

2.3 Denmark

At this moment KMS, Denmark (KMS, 2005) maintains a 1:10k dataset (which is currently being converted in an Oracle database) with an update cycle of five years. A dataset at scale 1:50k is currently automatically being derived from the 1:10k dataset using Lamp52/Clarity and various self written algorithms. The output from this automated process is manually checked and edited afterwards. The generalisation of buildings is currently being improved by self developed algorithms in Clarity. A 1:100k dataset is maintained independently from the other datasets since this map has a six year update cycle. First focus of KMS is to mature the 1:10k to 1:50k generalisation. The next focus will be on the generalisation from 1:10k to the other scales, based on the GeoDB concept (geo-database concept, see figure 2). KMS distinguishes between generic derived datasets (geoDBs) and cartographic output (maps, products/services). It should be noted that the geometry of objects is only changed when generalising the derived datasets. In principle converting from derived datasets to products/services never changes the geometry of the objects. Therefore the products and services are various symbolisations of the same datasets.

![Figure 2: GeoDB concept in KMS](image-url)
The role of KMS is currently changing from a national mapping agency to a governmental geospatial infrastructure co-ordinating organisation. Consequently main policy of KMS is that core datasets must be used as a basis for the digital infrastructure in Denmark, meaning that it is important to keep a unique key on objects. Cartographic output is no longer seen as the main goal. Current focus is on geospatial co-ordination, dataset specifications for different scales and purposes, quality control and database management.

2.4 France

IGN, France (IGN, 2005b) maintains two basic topographic datasets: BDTopo and BDCarto (apart from road datasets and DTMs). IGN is currently working on finalising BDTopo, which is a topographic 2.5D vector map with 1 meter accuracy (~1:10k), maintained in Geoconcept (one database for every map sheet). In 2007 BDTopo will cover the whole of France. BDCarto is a geographic vector database, with 10 meters accuracy (~1:50k) and already available for the whole of France. BDCarto is maintained to produce cartographic representation at scales 1:100k, 1:120k and 1:250k and is maintained in an Oracle database combined with ArcInfo. The first edition was digitised from the old 1:50k map. The semantic resolution between BDTopo and BDCarto differs, as well as the capture process. BDTopo is captured by stereo-restitution from aerial pictures, whereas in BDCarto a selection of features was captured based on existing 1:50k map sheets. The semantic resolution of BDCarto is not as good as the semantic resolution of a traditional 1:50k map, as for instance buildings have not been captured. Therefore the 1:50k map is not produced from BDCarto (the semantic resolution is not good enough).

IGN has a special research team at its disposal for long-term strategic research in order to support the production line. Research on generalisation started in 1992 and has yielded a lot of results in the areas of developments of platforms, generalisation algorithms, space characterisation, generalisation models, knowledge acquisition and evaluation of the generalised result. The research results are tested in projects, and if these projects achieve good results the research results are implemented in the production line. In IGN most generalisation that takes place is cartographic generalisation.

Cartographic generalisation is performed fully automatically from BDTopo to 1:25k map sheets using self developed software in DataDraw. Self developed software in research platform was used to derive cartographic representations at scales 1:100k and 1:120k once from BDCarto (one database per map sheet), with some minor generalisation activity (only displacement of highways). Once the dataset, underlying the 1:100k and 1:120k maps, was finished for the whole of France (2004), Lamps2/Agent technology was used to produce an updated version in one single database by only generalising the updates. The updated version was finished in February, 2005. The first edition of the cartographic representation at scale 1:250k was generated from BDCarto using DataDraw functionality (no generalisation) with road generalisation using Galbe (in house generalisation software). For the 1:250k, one database per sheet is maintained and updates are performed interactively in DataDraw. ArcInfo, with some human interactivity (selection) is used for database generalisation to produce road maps at scales 1:500k and 1:120K, as well as Euro Regional Map (ERM) and Euro Global Map (EGM) fully automatically from BDCarto. The 1:500k road map and the ERM are derived each year. The 1:120k road map and the EGM were derived once and are updated each year based on BDCarto. The updates are performed interactively. In a project environment (not yet in production) the generalisation of updates from BDTopo to BDCarto is tested. No link is maintained between BDTopo and BDCarto, neither between the derived datasets. At this moment it is studied how to maintain BDTopo and BDCarto in a multi-representation environment, while automatically propagating the updates (see figure 3).

![Diagram of possible future generalisation workflow in IGN, France](image-url)
2.5 Germany

Current generalisation workflow in Germany is illustrated in figure 4. Datasets at 1:250k and 1:1000k are maintained for whole Germany by the Bundesamt für Kartographie und Geodäsie (Federal Agency for Cartography and Geodesy (BKG, 2005)). The first edition of these datasets were produced by digitising the original maps (JOG250 and IWK1000), while now these datasets are updated through the generalisation of updates as indicated by the individual mapping agencies. Every province (Bundesland) in Germany has its own mapping agency responsible for topographic datasets and maps at scales 1:10k to 1:100k. Common goals have been identified for all mapping agencies in Germany. This has led to specifications for a Base-DLM and a DLM50 for the whole of Germany in the ATKIS project (Authoritative Topographic-Cartographic Information System) (ATKIS, 2005; Birth, 2003).

![Figure 4: Datasets in the ATKIS data model](image)

At this moment the Base-DLM is filled with data by digitising already available maps at scale 1:5k or 1:10k, while the content is based on maps of 1:25k, except for buildings and embankments. The feature classes and attributes are encoded corresponding to the ATKIS feature coding catalogue. The Base-DLM is used for topographic map production at scales 1:10k and 1:25k and as the basis for all other official datasets, as well as for external private customers. Map presentations in the scale range 1:10k and 1:25k can be generated semi-automatically from the Base-DLM using symbolisation. The ATKIS-DLM50 is used to produce (digital) maps at scales 1:50k ((D)TK50) and 1:100k ((D)TK100). It is also the basis for other official data in middle scale. Transition rules have been defined to translate between the Base-DLM and DLM50, see for example figure 5.

![Figure 5: Example of transition rules between base-DLM and DLM50.](image)

Although the Bundesländer have defined common goals concerning the Base-DLM and DLM50, at this moment three approaches are followed to obtain the DLM50 and DTK50 from the Base-DLM: AdV project (followed by seven Bundesländer, among which North Rhine Westphalia and Baden-Württemberg), the LGN application (followed by eight Bundesländer, among which Lower Saxony) and self-developed application of Bavaria. Bavaria did not participate in the workshop in Enschede, but its approach is very similar to the LGN application. After the DLM50 has been obtained once, only updates will be generalised.
The AdV project (Arbeitsgemeinschaft der Vermessungsverwaltungen, (ADV, 2005)) distinguishes between model generalisation and cartographic generalisation (see figure 6). The model generalisation is already in production and is performed 100% automatically in one step for a whole land. It should be noted that the topographic geometry of the Base-DLM is maintained in the produced DLM50, i.e. no displacements occur. Explicit references are built and stored between objects in the Base-DLM and in the DLM50. The cartographic generalisation is in development and aims to produce Digital Topographic Maps (DTKs) from the DLMs. This will be carried out by a mix of commercial software and self-developed software. It is expected to have still 30% to 40% human interaction.

![Diagram of generalisation workflow](image)

Figure 6: Future generalisation workflow in the AdV project

The LGN approach (LGN, 2005; Woldtke, 2004) does not distinguish between a topographic dataset and a cartographic dataset. Only one dataset is produced (called the DLM50) based on both model and cartographic generalisation. In this process CHANGE is used for building generalisation from scale 1:10k to 1:25k and from 1:10k to 1:50k. Model generalisation is based on automated self-developed algorithms with interactive assistance in complex situations based on SICAD/open software. The LGN algorithms are similar to those in the AdV project: selection, elimination, aggregation, merging, reclassification, typification, change of geometry type (e.g. from area to point representation), point reduction (Douglas-Peucker) and smoothing. The model generalisation is performed in spatial units. The cartographic generalisation is performed by interactive and semi-automated cartographic revision using self-developed SICAD/open tools. The result is a vector dataset with cartographic accuracy (compare with the AdV project). The Base-DML and DLM50 are stored independently and updates are performed in parallel for both datasets (in contrast to the AdV project). Future policy within the LGN approach is on improving running automatic and interactive generalisation processes based on the current ATKIS data model and SICAD/open capabilities, as well as on migration of all datasets and methods to the new ATKIS data model by 2007 based on ESRI/AED-SICAD software tools and Oracle database (Christoffers, 2005). Future production will integrate DLM- and DTK-updating within one working session, which means: updating of Base-DLM will include cartographic revision of DTK10 and DTK25. Next goal starting in 2008 will be the (semi-)automatic derivation of a DTK100 out of DLM50. Generalisation methods will have to be tailored to this new challenge.

### 2.6 Great Britain

The Ordnance Survey of Great Britain (OSGB, 2005) has a reference database, OS MasterMap, which is a seamless topographic database that covers the whole of Great Britain. The data have been collected at 1:1.25K in urban areas, 1:2.5K in rural areas, and 1:10k in mountain and moorland areas. The reference database is used to produce and deliver paper maps and digital products at scales 1:10k, 1:25k, 1:50k and 1:250k. OSGB makes extensive use of ESRI software, but uses also other software for specific purposes (Laser-Scan, Intergraph, MapInfo, etc.). Nowadays generalisation occurs in most of the production processes in OSGB. At the moment, different production lines use different software, with different custom tools. But most of the generalisation is done manually or by triggering simple tools interac-
tively. Consequently most of current generalisation requires human intervention. At the moment the different datasets are maintained independently from each other. The long-term goal is to update only the reference database and propagate the updates to the derived datasets, as automatically as possible.

As IGN France, also OSGB has a separate research team. The research on generalisation aims at building a system that can generate dedicated generalisation applications to produce all current and future products from a single information-rich database, while maintaining relationships between the different datasets. The research development platform is based on Clarity (from Laser-Scan) and extended to fit OSGB’s needs. The research plan contains four phases:

- build generalisation applications to produce maps automatically from OSGB reference data. The maps targeted are the existing OSGB Map series 1:25k, 1:50k and 1:250k, in order to facilitate the evaluation of the automatic results by comparing with existing maps. This phase will involve the development of generic algorithms and spatial analysis tools that will be reusable for other types of generalisation and designed for possible implementation on other production platforms such as ArcGIS.
- study how to build automatically the three generalisation applications mentioned above, from their specifications and the generic components (generalisation algorithms, spatial analysis tools).
- generate new generalisation applications from specifications of new products and the generic components of the system; this requires a lot of effort in formalising the product specifications, as well as controlling their integrity.
- generate new generalisation applications as in phase three, allowing the user to combine OSGB data with its own. This requires working on semantics and ontologies to allow the system to understand external data.

The research results on generalisation that were achieved so far are:

- strategy for generalisation has been defined
- development environment has been set up (based on Clarity)
- most research focus has been on generalising the 1:50k map from the OSGB MasterMap. The 1:50k project has been delivered successfully (development of spatial analysis tools, of generalisation solutions for building generalisation in rural and urban environment and of tools for collapsing dual carriageways), see also Revell et al, 2005

It is anticipated that the knowledge acquired and possibly the prototypes developed, will be gradually reused to improve the current map production flow lines.

2.7 The Netherlands

The Dutch Topografische Dienst (TDK, 2005) maintains datasets at scales: 1:10k, 1:50k, 1:100k, 1:250k, 1:500k and 1:1000k. The first editions of these datasets were produced by digitising the old, analogue maps. All scales are maintained in vector format and cartographic output can be generated automatically based on the vector datasets. Generalisation of updates takes place from 1:10k to 1:50k. The 1:100k is updated separately, since it has a different update cycle than the larger scale datasets. From the 1:100k the 1:250k is generalised, and from 1:250k the 1:500k is generalised. The generalisation is implemented in a customised interface with functions implemented on top of MicroStation. The generalisation process is currently an interactive process in which DGN files (MicroStation design files) are generalised according to rules described in the generalisation handbook (Topografische Dienst 1996/1998).

TDK has renewed the data model for the 1:10k dataset (Top10NL, see Bakker and Kolk, 2003) in order to meet modern requirements for geo-information: object oriented with unique coding, more attributes possible than in old dataset, indication of changes instead of the delivery of full updates, historical versioning, meta information at object level, separation between DLM and DCM, seamless database and based on open standards. The Top10NL for the whole Netherlands will be completed in 2005.

The first focus concerning generalisation will be the design of model generalisation to produce the 1:50k dataset based on Top10NL according to new specifications of the 1:50k dataset. In the meantime it will be studied how and with what software the generalisation from 1:10k to 1:50k can be implemented. After these steps have been taken generalisation to the smaller scale datasets will be studied. Generalisation processes in the future will be based on the DLMs. The DCMs will be generated by symbolisation of the DLMs.

2.8 Ireland

Ordnance Survey Ireland (OSI, 2005) maintains one large scale topographic dataset for which the scale varies (1:1k in urban areas, 1:2.5k in non-urban areas and 1:5k in remote areas). A smaller scale 1:10k cartographic product has been produced from which all small-scale digital and paper products are derived. This small-scale dataset is maintained separately from the large-scale topographic dataset. The current large-scale database of Ireland is limited in serving the growing needs of today’s geo-information society, in that the current dataset is map and sheet based; it has a proprietary data format; it is not ISO standard compliant; multi feature coding is used instead of feature attribution; it is not fully polygonised; and history and unique identifiers are only maintained at mapsheet level. First focus of OSI is on restructuring the traditional large-scale dataset into a single source large-scale database to overcome the current limitations. The first step (currently underway) is to convert the already available data into an open database format (not to ISO standards) in an Oracle 10g database which is still a sheet-based dataset and which has basically the same structure.
as the old dataset, with some improvements (limited attribution possible, history at feature level and unique identifiers for the whole dataset). The second step will be redefining the data model, introducing ISO standards, integration of the different scales of the reference dataset (1:1k, 1.2.5K and 1:5k), full connectivity and attribution. This will finally lead to a PRIME database. At this moment the different specifications are being reviewed and compared in order to indicate differences and similarities. For example in the current dataset it occurs that 1:1k data contains less information for some themes than the 1:2.5k and 1:5k data. In the future both a topographic database (PRIME) and a core cartographic database will be maintained. The cartographic database will be derived as automatically as possible from the PRIME database (probably by the end of 2006). Generalisation will only become a focus after the restructuring of the database has been completed, which will be by the end of 2006. Then it will be studied how, and with what software, to produce a small-scale dataset based on the large-scale dataset. The reason to also have the availability of a small-scale dataset is to meet customer’s requirements to have a less-detailed and consequently cheaper dataset. It is expected that generalisation tools will be implemented as part of the production line not sooner than 2007.

2.9 Sweden

Apart from a property map with topographic information at scale 1:10k, the Lantmäteriet in Sweden (Lantmäteriet, 2005) maintains topographic maps at various scales:

• a middle scale digital topographic map at scale 1:50k, which covers 2/3 of the country
• a 1:100k digital topographic map, with national coverage (100%)
• specific information for the mountainous area at scale 1:100k (such as hiking trails, mountain huts etc)
• (digital) topographic maps at scale 1:250k and 1:100k

All datasets are maintained as vector databases and are partly also used for GIS purposes. In the current production line the smaller scale datasets are derived from a larger scale dataset. The generalisation is performed by a step by step approach using mainly interactive methods by projecting the large scale dataset as background while generalisation is done by manual digitising on the screen. Some processes such as selection and filtering are performed using basic GIS functionalities in ArcGIS as well as Safe software in FME. Updates for the different datasets are performed independently from each other. Some themes are generalised with a full automatic selection and filtering process, e.g. simplified roads from 1:100k to 1:250k and administrative boundaries from 1:250k to 1:100k.

The current data models have been defined for different purposes, consequently it may occur that some objects have richer data models at the smaller scales. A core dataset at scale 1:10k with a renewed data model (simplified version in mountainous areas) was completed in 2004. In the future the databases at all different scales must be based on the same data model to make it possible to derive databases at smaller scales from the basic core dataset. The future system should support the renewed data models, history management, the indication of only those changes that are relevant for a certain scale and continuous updates as well as periodical updates. The future system requires also some organisational changes. First of all a move is required from building up new databases to updating of databases. Secondly, the new approach requires new skills which may not be available within traditional staff.

2.10 Switzerland

At this moment swisstopo (swisstopo, 2005) maintains map based landscape models at scales 1:25k (VECTOR25) and 1:200k (VECTOR200) for GIS analysis. Cartographic datasets are maintained at scales 1:25k, 1:50k, 1:100k, 1:200k, 1:500k and 1:1000k. In the current process the generalisation, which is performed to propagate updates, requires 100% human interaction. Swisstopo is at the beginning of a major change in which the whole data management including generalisation will be reconsidered. The goal of the change is having a multiple representation system containing one reference database and DCMs at different scales with automatic update propagation using generalisation. The new system is expected to be in production in 2008.

First a reference dataset has been defined that is currently evaluated by swisstopo. This 2.5D reference dataset (Topographic Landscape Model) will be the base model for all topographic products such as maps and DCMs at different scales. The accuracy of the TLM is between 1m and 5m (1:10k for roads and buildings and 1:25k for other objects). DCM25 and DCM50 will be generalised directly from the TLM as automatically as possible. DCM100 will be derived from DCM50, also as automatically as possible.

For the smaller scales first a VECTOR200 dataset will be generalised based on the TLM. The DCMs 1:200k to 1:1000K will then be generalised from VECTOR200. Generalisation will focus on updates only. It is expected that still a lot of interactive enhancement will be needed after the automatic generalisation process in order to refine the results. In the whole process the relationships between objects in the TLM and the DCMs at different scales will be maintained. Major steps that would support generalisation within swisstopo would be the generalisation of contourlines and a relevance check during updates, since not all updates in the source model are relevant for the smaller scales. For example a geometry change of 5m is important for the base model to assure 1m accuracy, but it is irrelevant for cartographic models. Objects should only be updated if there is a gain of information.
3. CONCLUSIONS ON GENERALISATION WITHIN NMA’S IN THE 21ST CENTURY

Generalisation in current practice From the descriptions above, it can be concluded that it is hard to give a general answer to the question what is the state-of-the-art on generalisation within NMA’s because of NMA specific characteristics (see also Stoter, 2005). For example all NMA’s distinguish between cartographic and database generalisation. However the difference is ambiguous. Is displacement of objects only appropriate when generalising DCMs (databases) or also for the DLMs (maps)? The aim of the displacement is to improve readability of the cartographic output, however also in the DLM one may not want different objects such as roads and buildings to overlap. The NMA’s who have chosen to displace the objects also in the DLM (Denmark, LGN approach in Germany, Catalonia) argue that if customers do not want displaced objects, they should use larger scale (which means more precise and accurate) datasets. Common to all selected NMA’s is that they are reconsidering their current production lines by the implementation of generalisation tools that work as automatically as possible. Policy on generalisation within current NMA’s can be divided into four main activities.

The first activity is the restructuring of current datasets into data models that meet the requirements of today’s geo-information society (i.e. data delivery within a geo-information infrastructure (GII), history, unique id’s, object oriented databases, using standards). This step has been taken, or will be taken in the nearby future, by all the described NMA’s.

The second activity is the creation of explicit links between data models, as well as between objects at different scales. The AdV project in Germany provides already the possibility to build and maintain references between different datasets. Catalonia has adjusted its data models at different scales in order to keep the coherence between semantics at different scales. The other NMA’s are or will remodel(ling) their data models in order to assure consistency. Apart from the AdV project, no NMA maintains the relationships between datasets at different scales at this moment.

The third activity is to decide on the conceptual architecture of generalisation within the renewed production line. The way that has been chosen by all selected NMA’s is to start with existing datasets in the renewed data models, while only generalising the updates. If the specific dataset did not yet exist (e.g. 1:50k in Denmark, 1:50k in Germany, 1:50k and 1:25k in Catalonia) the dataset is generalised once, after which the updates are generalised. Another decision concerning the architecture is whether to follow the ladder or star approach (see EuroGeographics, 2005). In the ladder approach (Belgium, Germany) one large scale dataset is maintained and the smaller scales are derived from the large scale dataset in steps. The alternative is the star approach in which every small scale dataset is generalised from one large scale dataset. Denmark, France, Switzerland and Catalonia have chosen for a mixture of both. OSGB, OSI, the Netherlands and Sweden have still to decide on the ladder or star approach.

The fourth activity is the actual implementation of automatic generalisation processes aiming at a) generalisation of updates, and b) automatic update propagation in which only relevant updates are taken into account. Dynamic generalisation as well as scale-less datasets are not considered as realistic. In traditional production lines some pre-processing (selection and filtering) as well as functions for specific situations is done automatically, but in general the generalisation requires a lot of human interaction. The NMA’s of Catalonia, Denmark, France, Germany and OSGB have already made fundamental steps in introducing automatic generalisation within a renewed production line, although the research results of OSGB were only implemented in prototypes. Catalonia: to produce 1:10k map and 1:25k map from the 1:5k database, and the 1:100k map from the 1:50k map, Denmark: to produce 1:50k dataset from 1:10k dataset, France: the generalisation of 1:25k maps from BDTopo and smaller scales from BDCarto and Germany: the model generalisation in the AdV project and the derivation of ATKIS-DLM50 in all projects. Automatic update propagation is only implemented in France, between the topographic database and cartographic database at the same scale. Belgium, the Netherlands, Ireland, Sweden and Switzerland are also faced with requirements for automatic generalisation in order to provide major improvements in the production line, but have not yet put many resources on the problem. The experiences in the selected NMA’s show that a lot of different software is used, often in combination with self-developed algorithms. A common approach of generalisation does not exist. However, four NMA’s are working on bundling their experiences and knowledge on generalisation. IGN Belgium, IGN France, KMS and OSGB work together in the MAGNET project, which is a user’s group of Clarity (Laser-Scan). The MAGNET group, which will be extended with other Clarity-users, meets twice a year aiming at exchanging experiences and algorithms in order to learn from each other.

Research needed to better support generalisation within NMA’s The second main question of this paper is what research breakthroughs are needed to better support generalisation within NMA’s. Discussions during the workshop led to the conclusion that research results have not always found a way to practice because of three main reasons. Firstly, NMA’s are not only dependent on research results but also on what results are implemented by vendors (including data enrichment, the generalisation itself as well as a quality check afterwards). It is difficult for vendors to supply software that provide a general solution, while taking the individual NMA demands into account. Secondly, a major
factor in the gap between research and usefulness in production is robustness requirements, which are much higher in practice than in research. Research may yield good results in prototypes. However, to actually implement the results into production lines the results should also be successful on very large, nation-wide base-datasets which do already exist and which contain data imperfections. Fault-tolerance and fail-safe behaviour are key characteristics needed in practice, which are of less importance in research environments. The last reason for the gap between research results and practice is the subjectivity of generalisation which is a consequence of the complexity of the generalisation process. If two cartographers are given the same generalisation rules, they will come to different results, because of the interdependencies between different (types of) objects and the mutual effect of generalisation rules: one rule may produce conflicts for other rules which makes the sequence in which the rules are applied extremely important. As a result exception situations in generalisation are not rare. It is not easy to formalise generalisation rules in such a way that they always give the same outcomes. The computer rule ‘if..then..else’ need to be extended with a ‘but’ possibility, which will not be easy to implement.

However, the workshop did come up with a list of topics which need more attention. These topics do not necessarily reflect the need for generalisation research but merely reflect what needs to be implemented in (commercial) software, which raises the question whether these breakthroughs are the responsibility of researchers or vendors. It is obvious that it remains very important to make research results available for the NMA community. The following functionalities were considered to be essential for generalisation within NMA’s in the coming years:

- A comprehensive and wide approach of generalisation functionality (on both geometry and attributes) that take the context (e.g. mountains, rural, urban) and relationships between objects into account
- Generalisation functions independent of the used GIS and underlying data model, but adaptable to specific data models that contain already existing datasets. This means that:
  - implicit data should be made explicit (semantics)
  - current databases need to be enriched
  - algorithms in software should work on different data models
  - knowledge of cartographers need to be formalised
- Support for Multi-Representation Databases: maintenance of links between derived and original dataset, automatic updating of derived datasets, relevance check during updates
- Generalisation of contour lines
- Generalisation of map names
- Objective methods to evaluate generalisation results

It should not be forgotten that these issues are defined by the outcomes of the workshop and are therefore practice orientated, aiming at solving generalisation problems in the nearby future within NMA’s. For science it may be a challenge to address more fundamental problems aiming at solutions for the long-term future.

An international approach to the generalisation problem may help individual NMA’s to decide on future steps in production line also based on experiences of other NMA’s. In this process emphasis within NMA’s should be on defining the results on not on defining processes and rules, which should be addressed by research and vendors. However, because of NMA specific characteristics it will require lots of (re)modelling and implementing work for every NMA to introduce automatic generalisation within their own, specific production line. Therefore it is necessary for NMA’s to put special resources on the generalisation problem in order to push the developments of generalisation within NMA’s in the 21st century.

ACKNOWLEDGEMENTS
I could never have written this paper without the active contribution of all participants of the workshop. Therefore I would like to thank Jan De Waele and Anne Féchir (Belgium), Nicolas Regnauld (Great Britain), Ernst Jäger (Lower Saxony), Sabine Urbanke (Baden-Württemberg), Ulrich Düren (North Rhine Westphalia), Bernard Farell and Colin Bray (Ireland), Anne Ruas and Cécile Duchêne (France), Nico Bakker and Jeroen de Vries (The Netherlands), Novit Kreiter (Switzerland), Inger Persson (Sweden), Marlene Meyer and Peter Rosenstand (Denmark), Maria Pla and Blanca Baella (Catalonia), Peter Woodsford (EuroSDR) and Menno-Jan Kraak (ITC). I am also very pleased with their remarks on previous versions of this paper.

BIBLIOGRAPHY


IGN, 2005, Institut Géographique National (Belgium), www.ign.be

IGN, 2005b, Institut Géographique National (France), www.ign.fr

KMS, 2005, Kort & Matrikelstyrelsen (Denmark), www.kms.dk

Lantmäteriet, 2005, www.lantmateriet.se


OSGB, 2005, Ordnance Survey Great Britain, www.orndancesurvey.co.uk/oswebsite/


Stoter, J.E., NMA specific characteristics regarding generalization, ICA workshop on generalization and multi representations, A Coruna, Spain, July 2005.


Topografische Dienst, 1996, Generalisatievoorschriften TDN

Topografische Dienst, 1998 Handboek generalisatie


**BIOGRAPHY**

Jantien Stoter (1971) graduated in Physical Geography at Utrecht University in 1995 before beginning her career as a GIS specialist with the District Water Board of Amsterdam (1995-1997). From 1997 till 1999 she worked as a GIS consultant at the Engineering Office of Holland Rail Consult where she applied GIS analyses to support the planning of large infrastructure projects. Stoter’s university career started in 1999 as an assistant professor in GIS applications, section GIS technology, in the Department of Geodesy, Delft University of Technology. In 2000 she started her PhD research on 3D Cadastre, which she defended successfully in September 2004. In February 2004, she received the prof. J.M. Tienstra research award for her work, which is given every five years by the Netherlands Geodetic Commission (NCG) in order to promote geodetic research in the Netherlands. Since April 2004 she holds the position of assistant professor at ITC, Enschede, the Netherlands. Her main research and education responsibilities are generalisation of geo-information and multi-scale databases.