COLLABORATING FOR BETTER ON-DEMAND MAPPING

Sandrine Balley and Nicolas Regnauld

Ordnance Survey, Southampton
sandrine.balley@gmail.com
nicolas.regnauld@ordancesurvey.co.uk

1. VISION

The amount of data available on the Web, often for free, is exploding. This is creating an immensely rich source of information; however, it is very difficult to make the best use of it. The richness comes from the possibility to combine different types of information and by analysing them together, to see new phenomenon emerge, which can lead to publishing new information.

Maps are certainly one of the most powerful ways of visually analysing correlations between different data, as long as there is a geolocation element attached to them. We therefore believe that the next big challenge in our field (of automatic generalisation, but it's much wider than that) is to bring the power of mapping to the finger tip of the user who needs to analyse a variety of data.

The role of data providers like Ordnance Survey will probably gradually switch from providing their data in the form of ready-made maps, to providing them in a way that allows a user or an application to select the themes and level of details required to bring the topographic context to their own data.

This constitutes a complex problematic that will only be addressed gradually, through studies that complement one another. This is why we think collaboration is necessary. The aim of this paper is to reflect on such collaboration. Section 2 positions our goal in relation to existing on-demand mapping studies. Section 3 suggests collaboration points and methods.

2. THE PROJECT

Ordnance Survey is setting up a multi-representation database that will be feeding every product of the portfolio through predefined derivation workflows. In parallel, Ordnance Survey's research department is investigating how this new set-up could deliver custom products automatically, through dynamic derivation workflows.

2.1 Our goal

On-demand mapping means the production of a map guided by user's requirements. We list the main steps of this process as follows:

1. Capture and interpretation of user's requirements into formal map specifications. These specifications must include the content of the map, as well as the way it will get displayed (for a map product) or structured (for a dataset product). If the content of the map/dataset includes some external data (supplied by the user), the specifications should also contain the

information that will allow the system to integrate it with its reference data (constraints of geometrical alignment for example).

2. Identification of relevant producer's data composing the topographic background
3. Content transformation, potentially including (order of steps may vary):
   - Integration of user's data
   - Data enrichment
   - Model generalisation
4. Data portrayal
5. Cartographic generalisation
6. Export, potentially including schema transformations

Our long-term goal is to achieve a full automation of the above process. Following the creation of target product specifications, the on-demand mapping system would design and then run a dynamic derivation workflow. Even if time response matters, we do not aim at on-the-fly derivation which we believe would narrow the range of possible customisation. This is obviously a very ambitious goal and we don't think it can be reached within a single project.

Our medium term goal is to pave the way for our long-term goal. We aim to design a conceptual framework, i.e. a high-level architecture and associated models, for on-demand mapping. This framework would enable to build an on-demand mapping architecture progressively, by ensuring that different contributions can be integrated.

2.2 Related works

A few on-demand mapping architectures have been designed or even prototyped. Some of them involve generalisation, like in the GiMoDig project (Sarjakoski & Sarjakoski, 2007) and the DURP Ondergronden project (Foerster, 2010). Other projects focus on the selection of the relevant themes and their portrayal, like the expert mapping system of (Forrest, 1999). These projects have demonstrated the feasibility of turning user's requirements into formal map specifications, and of using these specifications in a mapping process. Their limitation is to involve predefined data sources and predefined derivation tools whose parameters are included in the map specifications. The next step will be to decouple target product specifications from available resources. Models are ready for it, thanks to research communities on semantic data modelling (Kuhn, 2003) (Gesbert, 2005), semantic web services (Klien, Lutz, & Kuhn, 2006) (Lemmens, 2006), generalisation web services (Neun & Burghardt, 2005) (Foerster, Burghardt, Neun, Regnauld, Swan, & Weibel, 2008) and generalisation processes orchestration (Touya, Duchêne, & Ruas, 2010).

Also relevant to our goal are research projects focusing on one specific step of the process described above. The step of collecting requirements is studied by Human-Computer Interaction (HCI) and usability research teams. As describing the needed product encompasses technical choices that the general public is not able to make (Harding, et al., 2009), the users are instead enquired about their planned activities and asked to react to proposals from the system. These proposals are based on graphical examples and easy-to-grasp concepts. This approach was used by (Hubert & Ruas, 2003) to tune a generalisation process and by (Christophe & Ruas, 2009) to design a customised legend. Real time content transformation is a much studied issue. (Cecconi 2003) (Bernier and Bédard 2007) focused on the generalisation step and showed that combining generalisation with multi-representation databases was improving and speeding up the process. A generic, ontology-driven data enrichment method has been proposed by (Lüscher, Burghardt, & Weibel, 2007). Data integration theory is much studied as well (Devogele, Parent, & Spaccapietra, 1998) (Fonseca, Clodoveu, &

Camara, 2003) (Abadie, 2009), but there is currently no real-time, user-oriented service enabling users to integrate (not just overlay) their data into a referential (Riedemann & Timm, 2003) (Grosso, 2009). As for the export step, (Balley, 2007) (Gnägi, Morf, & Staub, 2006) (Lehto, 2007) (Schade, 2009) proposed schema transformation methods for users needing not only a map to display but also a dataset to process.

3. COLLABORATING FOR BETTER ON-DEMAND MAPPING

The previous sections introduced our vision on on-demand mapping and reviewed existing contributions. This section describes our early work and explores why, on what, and how we could collaborate.

3.1 Early exploration

In the first stage of our project, we identified the different kinds of knowledge required for on-demand mapping, reviewed related research, sketched a high-level architecture, and elaborated two use cases. Figure 1 schematises our vision of an on-demand mapping system with its two engines (in ovals) and resources (in rounded rectangles). We soon realised the necessity of finding partners and setting up collaborations before getting more in-depth into the system modelling.

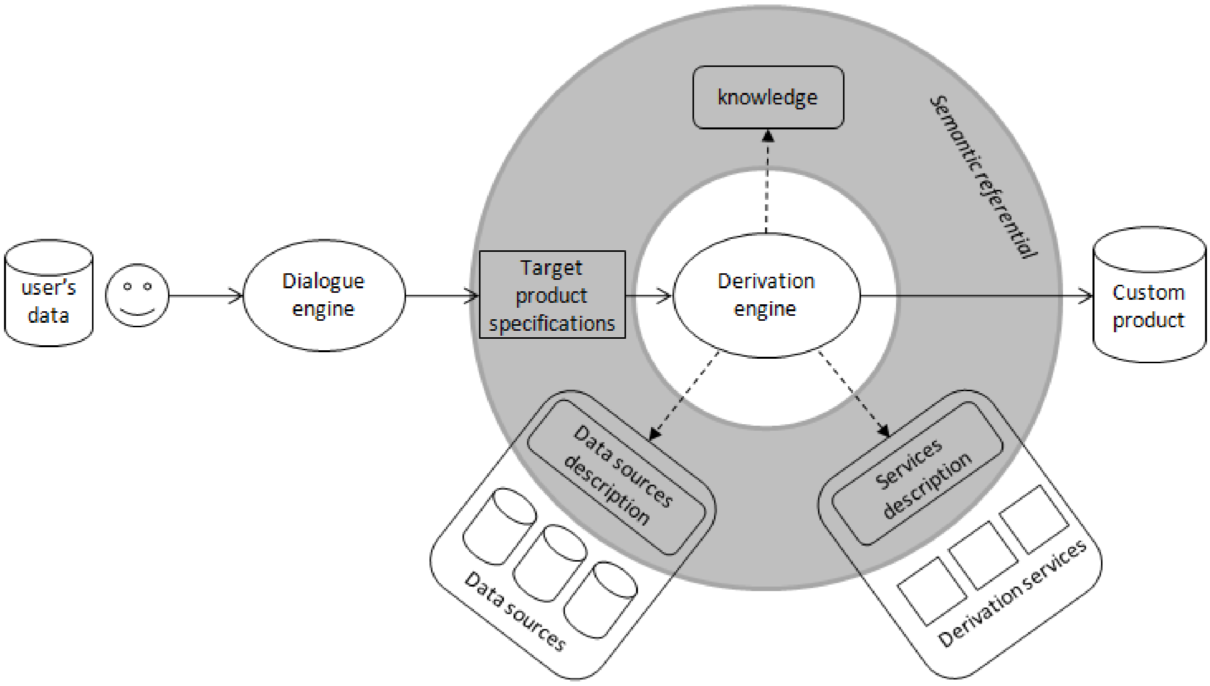


Figure 1: Overview of the on-demand mapping system. Plain arrows summarise the on-demand mapping process, from user's inputs to the custom product, with the target product specifications at the middle. To carry out the derivation process, the engine exploits (dashed arrows) resources represented in rounded rectangles.

3.2 Why collaborate?

The aim of the collaboration is not to build an on demand mapping system, but to define a framework that people can use to build such a system, by reusing existing components and adding new ones from their own expertise.

With this aim in mind, the first reason to collaborate is that on-demand mapping encompasses several research domains and thus would benefit from the expertise of several research teams. The second reason relates to interoperability requirements. On-demand mapping systems will probably be designed as distributed architectures involving web services for data access and generalisation, portrayal and schema transformation. Standard specifications (Web Feature Service, Web Processing Service and their profiles, Style Layer Descriptor) are available for some of these services (Foerster & Stoter, 2006), but some others still have to be specified through a community process. These services should be discovered, selected and chained dynamically by on-demand mapping systems, which requires semantic descriptions based on an agreed description model and agreed concepts.

3.3 Collaboration points

There are different and complementary collaboration approaches. The first approach is to divide the work along the different research challenges. The second approach is to agree on models. Lastly, to instantiate these models, we propose to define a common semantic referential.

3.3.1 Separating research challenges

The process of deriving a custom product, outlined in section 2.1, presents research challenges that could be addressed separately.

The collection of the user's requirements presents a HCI challenge. As requirements may concern several facets of the product, several interaction modes will probably be relevant. A flexible path should allow the user to describe their need in the order they chose, or even to give minimal indications and refine them later, when required by the system for building the workflow.

Reasoning capabilities, based on usability knowledge, must be implemented in order to instantiate the formal product specifications according to the collected requirements (Forrest, 1999) (Sarjakoski & Sarjakoski, 2007) (Bucher, Buard, Jolivet, & Ruas, 2007). The specification of custom-made, optimal legend (Christophe & Ruas, 2009) can be isolated as a specific challenge requiring knowledge about semiotics.

Data integration is another challenge. Formal metadata need to be produced, for linking precisely the data sources (reference data from the data provider and user's data) to the real world concepts used to specify the target product, and also used to describe the processes available to transform the data (data integration, generalisation, etc.).

The biggest challenge is the dynamic workflow design, enabling to automatically plan the steps of content transformation, portrayal, cartographic generalisation and export We need a set of chainable abstract operations and probably a set of predefined high-level tasks based on them (Bucher, 2009). We also need a large amount of formalised knowledge (procedural knowledge and domain knowledge about maps, legends, datasets and data schemas) helping the system to sequence operations and specify their variables. For the execution of the workflow, the sequence of abstract operations must be translated into a chain of web services, discovered using their semantic description.

3.3.2 Sharing models

In order to make the modules interoperate, it is necessary to adopt common models for the core components of the architecture (represented on figure 1) and for the system resources.

Target product specifications model

The target product specifications are a key element of the system, as they act as the link between two very different domains. Specifications are the recipient for the task focusing on collecting the user requirements, and they are an input to the derivation engine. A common formal specification model must be defined. It must be independent from available data sources and derivation tools, machine-readable and rich enough to enable an engine to build a custom derivation workflow.

Data sources description model

The system must select the most relevant data source and identify what transformations are required to make it compliant with the target product specifications. For that purpose, data sources must be described not only in terms of real world concepts (roads, rivers) but also in terms of how the real world entities are represented (which rules drove the selection of real world entities? How was the entity observed and geometrically modelled? Which rules drove the encoding of entities properties into feature attributes?). Such descriptions can be formalised through a specification model linking ontologies of real world concepts with databases schemas as proposed by (Gesbert, 2005). This cannot be achieved by current standard models for data specifications (ISO, 2005) and feature concept dictionaries (ISO, 2009).

Derivation tools description model

In order to build the derivation workflow, the system must have a set of abstract operations available with their functionalities (e.g. simplification, reclassification), roles described in terms of real world concepts (e.g. building) and GIS constructs (e.g. size threshold, feature type), pre- and post-conditions. Research works about service chaining (Lemmens, 2006) or process descriptions (Bucher & Jolivet, 2008) have provided such models, which could be instantiated with concepts from the semantic referential. In order to run the derivation workflow with available derivation services, the system then needs to access service descriptions matching the previous operations descriptions. OWL-S or semantic-annotated WSDL might be suitable models.

Knowledge representation model

The knowledge required by the system can be classified into:

- Product design knowledge to create relevant target product specifications,
- Cartographic and GIS knowledge to parameter each derivation operation,
- Procedural knowledge to guide the workflow design.

This knowledge can take the form of rules, constraints, user profiles and specification templates. In the generalisation domain, the content and structure of cartographic constraints used by different data providers have been compared and harmonised (Burghardt, Schmid, & Stoter, 2007). Although it is just an option (knowledge is a trademark of each data provider and doesn't necessarily have to be shared), doing the same for all domains of on-demand mapping knowledge would enable to exchange and test knowledge in different map derivation systems.

3.3.3 Building a common semantic referential

The models described in section 3.3.2 are semantic models, i.e. they rely on concepts. This enables to:

- Connect the on-demand mapping system with the user need, which is expressed through real world concepts,
- Ensure the system stays generic and can potentially handle new resource (data or derivation service) described through the agreed model

To instantiate these models, we propose to define a semantic referential (Kuhn, 2003), i.e. a repository of all allowed concepts. Organised into taxonomies or ontologies, these concepts will constitute the base vocabulary of the target product specifications, of the system knowledge and of the resources descriptions. This semantic referential will also be a collaboration tool, helping partners to understand each other. Three sets of concepts will be defined: real world geographic concepts (such as road, church or width), GIS concepts (such as feature, threshold or stroke) and operations (such as simplification, buffer or amalgamation).

We are aware of the difficulty to agree on geographic concepts, and indeed we don't seek an absolute consensus on the terms definitions (as far as we don't propose trans-border products). What we need is to make sure that an operation described as manipulating a geo concept like "roads" is relevant to use on different data sources representing roads. We believe this requires a moderate granularity level in term definitions, allowing us to ignore slight nuances at least during the stages of investigation and prototyping. That is why we favour ready definitions, e.g. from Inspire specifications.

GIS concepts and operations are more easily agreed on. To define them we can rely on OGC definitions and existing operations taxonomies (Mustière, 2001) (Foerster, 2010) (Regnauld & McMaster, 2007) (Lehto, 2007).

3.4 How to collaborate?

From informal exchanges to funded projects, collaboration means still have to be defined.

We are already collaborating with the Manchester Metropolitan University via the PhD project of Nicholas Mark Gould on dynamic workflow design, which started in 2011.

We are also collaborating with the Cogit laboratory of IGN-France, where several aspects of on-demand mapping are studied. We have started building a semantic referential, limited to the context of one of our use cases. This is being done through a wiki illustrated on Figure 2.

4. CONCLUSIONS

Allowing non expert users to design and build the map that suits their needs, combining the data of their choice with reference data, is a very big challenge. The system will have to provide the expertise from different domains to the user: automatic data integration, automatic generalisation, map/legend design, capturing user needs.

In order to take on the challenge, we believe that the best way is to follow a progressive path, where we focus on allowing the integration of new components for different parts of the system. In this way, libraries of tools, data, profiles, knowledge can be built and enriched. These can then be used to design and build automated map derivation systems that get more and more powerful as the libraries get richer.

For this to happen, we need identify the key components of the system that require formal descriptions or standards. Identifying these and formalising them should be done in a collaborative way, to avoid creating an "in house" solution, which would have little chance of being widely adopted and evolve.

The first step of the collaboration could be to form a consortium that maintains a Website with information and links to resources available. One of the resources, which should probably be the first task of the consortium, would be a vocabulary to describe all the resources of the system (data, tools and knowledge).


page | discussion | edit | history | delete | move | protect | watch

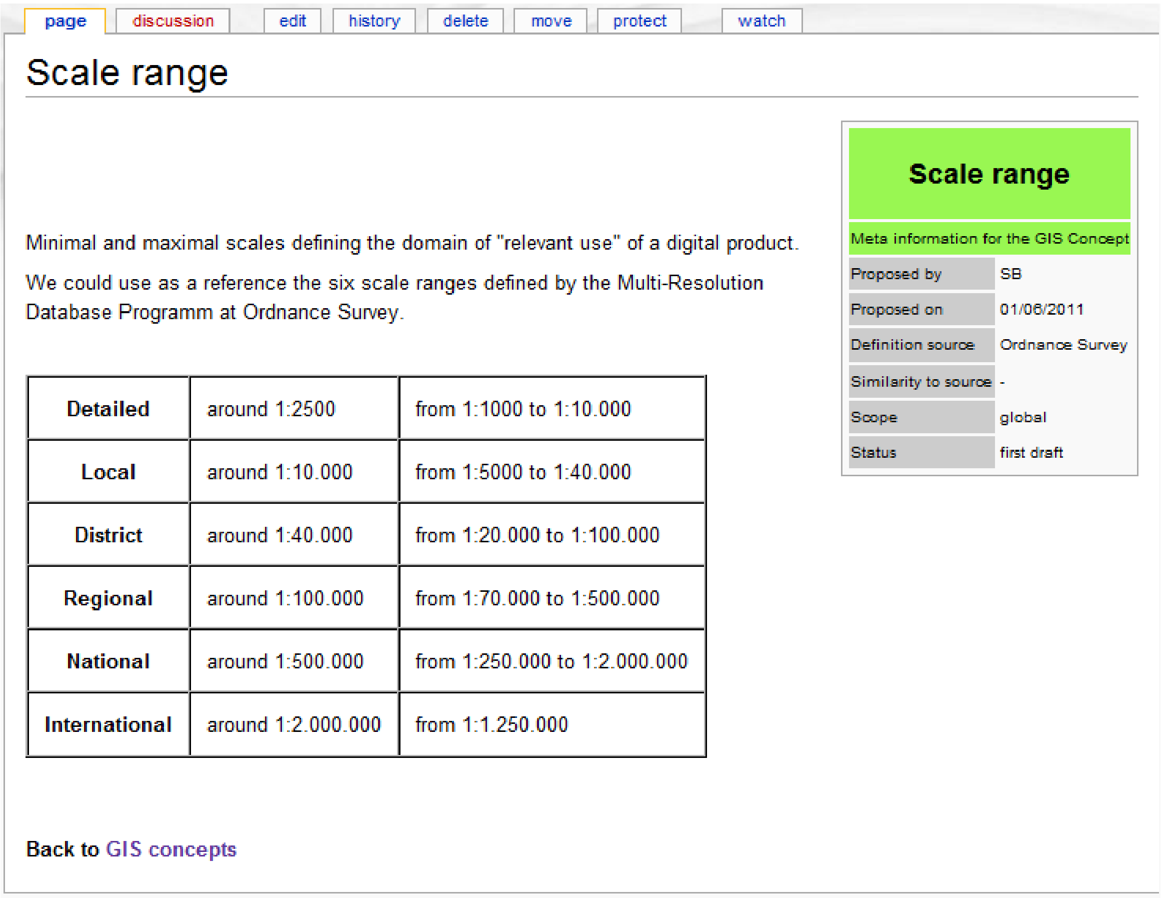
Motorway

Definition [edit]

A major classified road, and associated approach road and slip roads, consisting of separate carriageways with separate flow directions, the carriageways are partitioned by barriers.

Back to Geo Concepts

Motorway	
Meta information for the Geo Concept	
Proposed by	SB
Proposed on	04/05/2011
Definition source	OS RWO Catalogue
Similarity to source	identical
Scope	first use case
Status	first draft


page | discussion | edit | history | delete | move | protect | watch

Scale range

Minimal and maximal scales defining the domain of "relevant use" of a digital product.

We could use as a reference the six scale ranges defined by the Multi-Resolution Database Programm at Ordnance Survey.

Detailed	around 1:2500	from 1:1000 to 1:10.000
Local	around 1:10.000	from 1:5000 to 1:40.000
District	around 1:40.000	from 1:20.000 to 1:100.000
Regional	around 1:100.000	from 1:70.000 to 1:500.000
National	around 1:500.000	from 1:250.000 to 1:2.000.000
International	around 1:2.000.000	from 1:1.250.000

Back to GIS concepts

Scale range	
Meta information for the GIS Concept	
Proposed by	SB
Proposed on	01/06/2011
Definition source	Ordnance Survey
Similarity to source	-
Scope	global
Status	first draft

Figure 2: The Motorway *Geo concept and and the* Scale range *GIS Concept defined in our wiki (draft definitions, not approved by IGN).*

REFERENCES

Abadie, N. (2009). Schema Matching Based on Attribute Values and Background. *12th AGILE International Conference on Geographic Information Science.* Hannover.

Balley, S. (2007). *Aide a la restructuration de donnees geographiques sur le Web - Vers la diffusion a la carte d'information geographique.* PhD Thesis, Université Paris Est - Marne la Vallee.

Bernier, E., & Bédard, Y. (2007). A data warehouse strategy for on-demand multiscale mapping. In W. Mackaness, A. Ruas, & L. Sarjakoski (Eds.), *Generalisation of Geographic Information: Cartographic Modelling and Application, édité par W.A. Mackaness, A. Ruas et L.T. Sarjakoski, 67-87. Elsevier, 2007.*

Bucher, B. (2009). *Vers la Diffusion en Ligne d'Information Géographique sur Mesure.* ASR thesis, Université Paris-Est Marne-la-Vallée.

Bucher, B., & Jolivet, L. (2008). Acquiring service oriented descriptions of GI processing software from experts. *11th AGILE International Conference on Geographic Information Science* . Girona.

Bucher, B., Buard, E., Jolivet, L., & Ruas, A. (2007). The need for Web legend services. In J. a. G.E.Taylor (Ed.), *7th International Symposium on Web and Wireless GIS.* Cardiff: Springer Lecture Note in Computer Science.

Burghardt, D., Schmid, S., & Stoter, J. (2007). Investigations on cartographic constraint formalisation. *Workshop of the ICA Commission on Generalisation and Multiple Representation.* Moscow.

Devogele, T., Parent, C., & Spaccapietra, S. (1998). On spatial database integration. *International journal of geographical information science , 12* (4), 335-352.

Cecconi, A. (2003). *Integration of generalisation and multi-scale databases for enhanced web mapping.* PhD Thesis, University of Zurich.

Christophe, S., & Ruas, A. (2009). A process to design creative legend on-demand. *11th International Cartographic Conference.* Santiago, chile.

Foerster, T. (2010). *Web-based architecture for on-demand maps - Integrating meaningful generalisation processing.* PhD Thesis, University of Twente.

Foerster, T., & Stoter, J. (2006). Establishing an OGC web processing service for generalization processes. *ICA workshop on generalization and multiple representation.* Portland, Oregon, USA.

Foerster, T., Burghardt, D., Neun, M., Regnauld, N., Swan, J., & Weibel, R. (2008). Towards an Interoperable Web Generalisation Services Framework – Current Work in Progress. *11th Workshop of the ICA Commission on Map Generalisation and Multiple Representation.* Montpellier.

Fonseca, F., Clodoveu, D., & Camara, G. (2003). Bridging Ontologies and Conceptual Schemas in Geographic Information Integration. *GeoInformatica , 7* (4), 355-378.

Forrest, D. (1999). Developing Rules for Map Design: A Functional Specificationfor a Cartographic-Design Expert System. *Cartographica , 36* (3).

Gesbert, N. (2005). *Étude de la formalisation des spécifications de bases de données géographiques en vue de leur intégration.* PhD Thesis, Université de Marne la Vallée.

Gnägi, H. R., Morf, A., & Staub, P. (2006). Semantic Interoperability through the Definition of Conceptual Model Transformations. *9th AGILE Conference on Geographic Information Science.* Visegrad.

Grosso, E. (2009). Data Integration GeoService:A First Proposed Approach Using Historical Geographic Data. In A. S. James D. Carswell (Ed.), *9th International Symposium on Web and Wireless GIS*, (pp. 103-119). Maynooth.

Harding, J., Sharples, S., Haklay, M., Burnett, G., Dadashi, Y., Forrest, D., et al. (2009). Usable geographic information – what does it mean to users? *AGI Geocommunity.*

Hubert, F., & Ruas, A. (2003). A method based on samples to capture user needs for generalisation. *6th ICA Workshop on Generalisation and Multiple Representation.* Paris.

ISO. (2009). *ISO-TC211 - 19123 Geographic information: Feature concept dictionaries and registers.* International Standard.

ISO. (2005). *ISO-TC211 - 19131 Geographic Information - Data Product Specification.* International Standard.

Klien, E., Lutz, M., & Kuhn, W. (2006). Ontology-based Discovery of Geographic Information Services – An Application in Disaster Management. *Computers, Environment and Urban Systems , 30* (1), 102-123.

Kuhn, W. (2003). Semantic Reference Systems. *International Journal of Geographical Information Science , 17* (5), 405-409.

Lehto, L. (2007). Schema Translations in a Web Service Based SDI. *10th AGILE International Conference on Geographic Information Science.* Guimarães.

Lemmens, R. (2006). *Semantic interoperability of distributed geo-services.* PhD Thesis, ITC.

Lüscher, P., Burghardt, D., & Weibel, R. (2007). Ontology-driven Enrichment of Spatial Databases. *10th ICA Workshop on Generalisation and Multiple Representation.* Moscow.

Mustière, S. (2001). *Apprentissage Supervisé pour la Généralisation Cartographique.* PhD Thesis, Université Paris 6.

Neun, M., & Burghardt, D. (2005). Web Services for an Open Generalisation Research Platform. *8th ICA Workshop on Generalisation and Multiple Representation.* A Coruña, Spain.

Regnauld, N., & McMaster, R. B. (2007). A synoptic View of Generalisation Operators. In W. A. Mackaness, A. Ruas, & L. T. Sarjakoski (Eds.), *Generalisation of Geographic Information: Cartographic Modelling and Applications* (pp. 37-66). Elsevier.

Riedemann, C., & Timm, C. (2003). Services for data integration. *Data Science Journal (Spatial Data Usability Special Section) , 2.*

Sarjakoski, T., & Sarjakoski, L. T. (2007). A Real-Time Generalisation and Map Adaptation Approach for Location-Based Services. In W. Mackannes, A. Ruas, & L. T. Sarjakoski (Eds.), *Generalisation of Geographic Information: Cartographic Modelling and Applications* (pp. 137-159). Elsevier - In behalf of the International Cartographic Association.

Schade, S. (2009). *Ontology-Driven Translation of Geospatial Data.* PhD Thesis, Universty of Muenster.

Touya, G., Duchêne, C., & Ruas, A. (2010). Collaborative Generalisation: Formalisation of Generalisation Knowledge to Orchestrate Different Cartographic Generalisation Processes. *GIScience.* Zurich.