Interoperability in Map Generalisation Research

Alistair Edwardes, Dirk Burghardt, Moritz Neun

Department of Geography, University Zurich
Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
{aje, burg, neun}@geo.unizh.ch

Abstract. There is a pressing need within the map generalisation research community to initiate processes for supporting cooperation and allowing techniques and results to be shared. On the one hand, increasing complexity of methods and concepts in this domain has made the issue of open access to research evermore necessary to ensure that new investigations can be undertaken at the research frontier with the minimum of ancillary effort. On the other, as the basis for conducting 'good science', researchers must be able to gain access to growing body of prior work, so that comparisons can be made and results evaluated. Realising such desires involves addressing a complex set of interoperability challenges at a number of different levels; technical, syntactic and semantic. The paper describes the experiences of the authors in tackling such issues. It first details the implementation of a web services technology for exposing generalisation research through an open interface. Then it discusses the need for formalized data structures to encode geographic representations and generalization concepts. Finally it considers the issue of semantic interoperability and examines how ontologies might be used to overcome these difficulties.

Keywords. Automated generalisation, common research platform, standards, semantics, ontology, urban generalisation, web services

1 Motivation: an open research platform for generalisation

The field of automated map generalisation has been a fertile area of research for over thirty years. In the past, much of the research was focussed around questions that could be explored largely independently of other researchers. However, as the field has matured there has been an increasing need to operate more effectively as an integrated community. One reason for this is that as a better understanding of the generalisation processes develops the complexity of solutions increases. Hence, researchers must spend more and more time gathering tools to reach the research frontier. In addition, this evolution results in body of core concepts developing as the most effective methods for stratifying the domain are discovered (e.g. subdivisions of operators and generalisation processes). It therefore becomes important to harmonise definitions for these concepts so they can be effectively re-used as a basis for exchange amongst researchers. In the interests of 'good science', it is also important to be able demonstrate new methods are improvements over what has already been achieved. This requires access to comparable methods for performing appraisals. Being able to share techniques and models is thus essential for the efficient working of the field.

Recently, there has been a growing movement towards the development of open systems to achieve these goals. These efforts are described in detail elsewhere (Edwardes et al, 2003; Burghardt et al, 2005). The current focus has been to adopt the mechanism of web services to support sharing and interoperation amongst researchers in a platform independent way. A prototype has been developed to achieve this goal technically and physically (Neun and Burghardt, 2005), however, what is lacking is the formalisation and specification of domain concepts that would take this to a more logical level.

This paper looks at how a web-services based open research platform might be extended in this way. It first, describes the state of the art for web services in generalisation research. Then, it discusses issues involved in enhancing interoperability by formalising generalisation concepts, considering which are the most important concepts in the generalisation domain. Finally, it considers the semantic barriers to interoperability and how ontologies might help by modelling generalisation organisations for specific types geographic phenomenon.

2 Web services to support map generalisation research

Web services are generic mechanisms for accessing computational objects and operations over the web. In the web services model (WSA, 2004) operations are presented through XML interfaces (SOAP, 2003) allowing them to be accessed in a platform independent way. Services are exposed using the "publish-find-bind" paradigm (UDDI, 2004). *Publishing* involves the service provider creating an interface describing the parameters of the operations and declaring the URL of where the service can be accessed, its *endpoint*. This is often done through an intermediary *registry* of web services. The *find* step is performed by a service consumer. They selected the desired service from those available in the registry. Using the interface description given there the consumer can then *bind* to an operation. This establishes communication with the service endpoint. In a generalisation service the service provider might be a researcher publishing and algorithm and the consumer a researcher wishing to use that algorithm in their work.

There are several advantages of using generalisation services in a collaborative and distributed research environment additional to those mentioned previously. First of all, the platform independence makes the development independent from the operating system and hardware used and allows parties to share without needing to expose or share their source code (particularly important for more commercial research institutions). Secondly, the service can be integrated in any software platforms, such as web browsers, GIS or map production software. Thirdly, specific algorithms for special computer architectures e.g. clusters, grids or other parallel processing systems can be offered. Lastly, the service can be accessed over the internet or locally. Figure 1. Illustrates the registry based WebGen architecture (Neun and Burghardt, 2005) for generalisation web services.

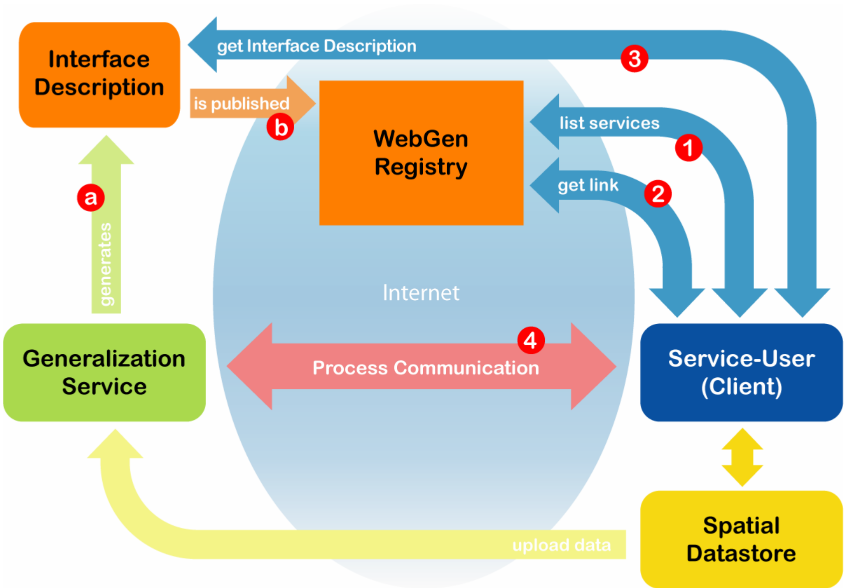

Figure 1 The WebGen architecture (Neun and Burghardt, 2005)

In Figure 1 the publish step is shown at *a* and *b*. *1, 2, 3* comprise the find step. Binding is performed in *4*.

The complexity of generalisation services can be separated (Figure 2) into three different levels of abstraction (Burghardt et al., 2005):

1. **Generalisation support service:** The main goal of these services is to support generalisation operations by making spatial information explicit. This

might be by representing common structural properties such as neighbourhoods (e.g. with buffers, topology or voronoi diagrams) or proximity relations (e.g. with Delaunay triangulations) which can then be exploited by other generalisation operations (Neun et al., 2004) and optionally stored in a Multi-Resolution Database (MRDB).

2. **Generalisation operator services**: These deliver the functionality of standalone generalisation operators. Examples are services for simplification, smoothing, aggregation, amalgamation, merging, collapse, refinement, exaggeration, enhancement and displacement. These generalisation operator services can be further subdivided for point, line and area objects and specialised depending on object classes.
3. **Generalisation process services**: These services offer functionality to support the control and orchestration of the generalisation process. Examples are services for automated orchestration, services for evaluation of generalisation results and agents (Lamy *et al*, 1999).

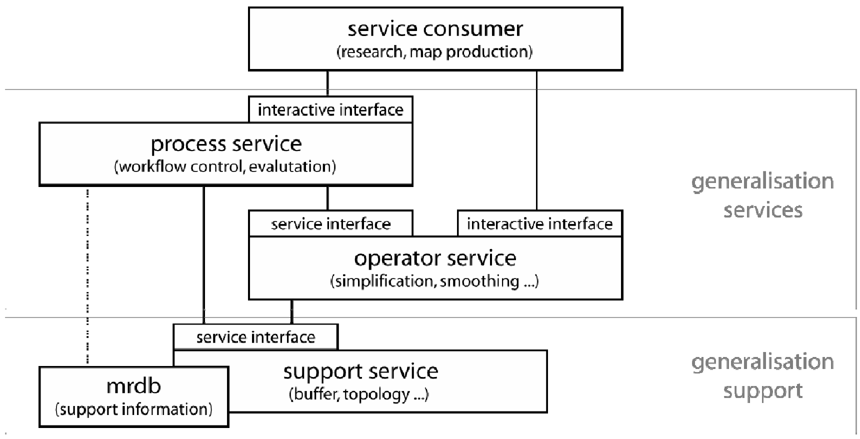

Figure 2 Framework for generalisation services.

Both server and client web service components have been implemented. On the server side a registry has been implemented using Java servlets. This offers a single access-point where all information can be found on which services are available, where they are located and what algorithms they offer. Service providers can use this to *publish* their operations. On the client side, components to realise two usage scenarios have been implemented. In the first case access to the operations is through a web application. Here, users *find* the service they want through a web interface to the registry, they then setup parameters using web forms related to the selected operations. *Binding* is performed by posting data to be processed as Shapefiles to the service. The advantage of this approach is that it only requires a web browser for access. The second scenario uses plug-ins for mapping software. Currently the only plug-in developed is for the open-source JUMP platform (JUMP 2005) though support for other platforms such as ESRI's ArcView® is planned. The advantage of the plug-in approach is that it allows close integration with mapping data and operations providing a transparent and seamless user experience. Figure 3 shows the JUMP plug-in with different implemented generalisation operations.

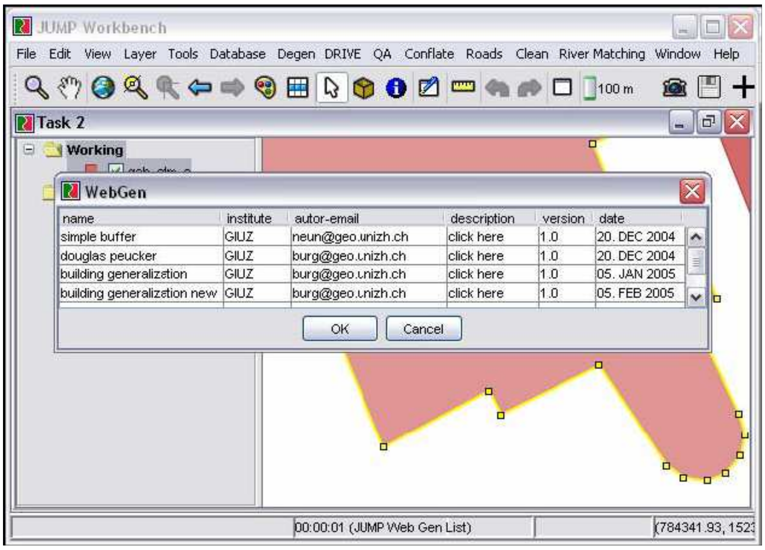

Figure 3 JUMP plug-in

3 Conceptualisation of the generalisation domain

Interoperability of domain concepts occurs at two levels; the syntactic and the semantic (Vckovski, 1998). The syntactic level deals with the encoding of geographic abstractions, for example as data types. The semantic level considers the inconsistencies (Shepard, 1991) amongst different users' representations of the world for different purposes.

3.1 Syntactic interoperability

Syntactic interoperability involves harmonising heterogeneities amongst different geographic data abstractions. In a web service this kind of interoperability is ensured by adherence to commonly agreed standards for data abstractions and encodings at the service interface. Table 1 shows how the WebGen service uses the standard data encoding GML in the interface definition of an operation.

```
<?xml version="1.0" encoding="UTF-8"?>
<webgen  xmlns:gml="http://www.opengis.net/gml"
xmlns:SOAP-ENC="http://schemas.xmlsoap.org/soap/encoding/">
<name>building simplification</name>
<method>buildingSimplification</method>
```

```
<endpoint>http://www.geo.unizh.ch:8080/neun/servlet/GenHandlerXML
</endpoint>
<description>This algorithm simplifies all buildings in a layer and
returns a layer containing the resulting buildings!</description>
<config>
 <layer>
  <schema>
   <attribute name="geom" type="GEOMETRY">
    <allowed>gml:Polygon</allowed>
    <allowed>gml:MultiPolygon</allowed>
   </attribute>
  </schema>
 </layer>
 <param name="min edge length” type="SOAP-ENC:double" >
  <description>Minimum Edge Length</description>
 </param>
</config>
</webgen>
```

Table 1 XML interface for a WebGen operation

The WebGen interface definition extends a standard WSDL (2001) interface for web services to allow geometric data encoded in the GML format to be passed across it. This extension is made by importing the GML namespace, shown in bold in Table 1.

Using a common abstraction for data at the geometric level provides the most basic level of integration required by generalisation operator services. It allows simple generalisation operators processing single features to be offered e.g. line and building simplification. In addition it allows generalisation operators to be integrated with the same abstractions used for more general GIS software, providing a link between technologies. However, more complex generalisation requires new formalisation of data types to be defined. This needs to be done both within the map generalisation community and between the community and other related fields, for example map production, computational geometry and spatial analysis. We can consider the different requirements for formalisation by looking at the different types of generalisation web service in turn.

3.1.1 Generalisation Support Services

These types of services realise representations that abstract aspects of the spatial or topological relations inherent within the data. Amongst the most important data types here are those that describe planar graphs, since these can be used as generic concepts to describe a variety of different data structures. GML3 (GML 2004) provides a standard set of primitive classes to encode these in GML as topological nodes, edges and faces. Regnauld (2005) presents a more comprehensive model respecting the specific requirements of graphs for generalisation. He shows how the model can be used as the basis for representing structures including Delaunay triangulations, Voronoi diagrams, transport graphs and minimum spanning trees. The class decomposition is shown in Figure 4.

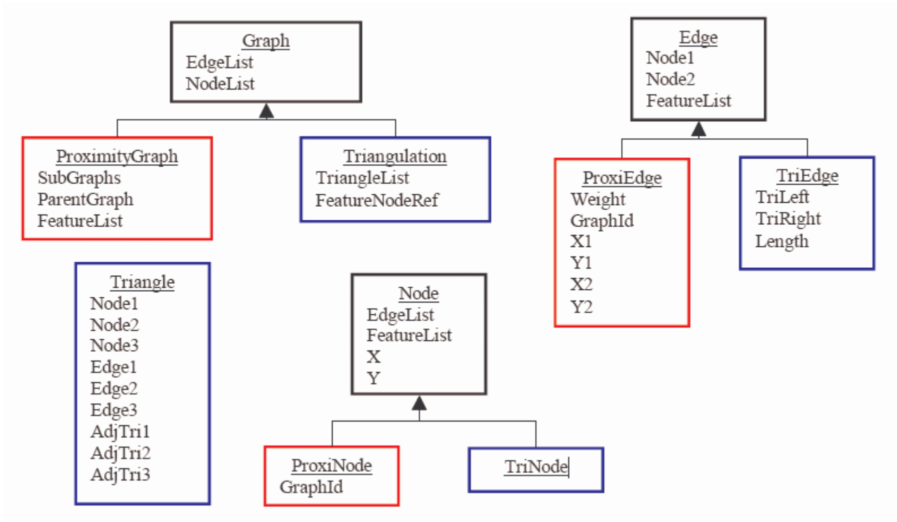

Figure 4 Classes for graphs (Regnauld, 2005)

3.1.2 Generalisation Operator Services

Various authors have suggested typologies for generalisation operators (Hake *et al*, 2001; McMaster and Shea, 1992; Bader *et al*, 1999). Whilst these have many similarities their discrepancies can cause difficulties in making comparisons or studying sequencing operations. A more standardised taxonomy would help resolve this. Table 2 describes the operators organised in the typology of Bader *et al* (1999) and used in the AGENT project (Lamy *et al*, 1999) that might be used in this regard.

Classification	Thematic Selection	
	Thematic Aggregation	
Simplification	Weeding	
	Unrestricted Simplification	
Collapse		
Enhancement	Geometric Enhancement	Enlargement
		Exaggeration (caricature)
	Semantic Enhancement	Smoothing
		Fractalization
		Rectification/ Squaring
Selection / Elimination	Selection	
	Elimination	
Displacement		
Aggregation	Amalgamation	Fusion
		Merge
	Combine	

Table 2 Typology of generalization operators (Bader *et al*, 1999)

3.1.3 Generalisation process services

These services provide access to complex and fully-automated generalisation procedures. Here, interaction with the service is *declarative*, allowing the user to state their intentions for a generalised map rather than simply prescribe operations that should be performed in a *procedural* manner. Allowing declarative statements requires a set of concepts that allow the user to describe their aims. Considerable success has been achieved using the constraint-based approach to accomplish this (Beard, 1991). Here the user asserts the constraints the final solution should satisfy and allows some form of automated procedure to search for a solution that best satisfies these. Constraints might include; the minimum width of a feature, the minimum separation distance between two features or the minimum size of a feature. Examples of automated constraint-satisfaction procedures are multi-agent systems (Ruas, 1999), least-squares adjustment (Harrie, 1999; Sester, 2000) and simulated annealing (Ware et al., 2003). The model of constraints of Ruas and Plazanet, (1996) is illustrated in Table 3.

Legibility Constraints	Minimum separation
	Granularity
	Minimum edge length
	Minimum width
Shape Constraints	Single object
	Set of objects
Spatial Constraints	Absolute position
	Relative displacement
	Topological relations
	Proximity relations
Semantic Constraints	Quantity
	Inter-classes quantity
	Function

Table 3 Typology of Constraints by Ruas and Plazanet, (1996)

Constraints are defined in unison with *measures* and *algorithms*. Measures are analysis tools that evaluate the degree to which a constraint is satisfied. Algorithms are generalisation operators that improve the status of a constraint. Formalised data types for each of these concepts are needed together with mechanisms to prioritise amongst them. One solution for this is the XMLSchema model suggested in Hardy *et al* (2003).

3.2 Semantic interoperability

The ultimate goal of generalisation is to allow map users to reason about geography at multiple levels of abstraction. Complex generalisation operations therefore seek to derive from source data new abstract geographic forms and correspondences amongst them in order to provide new views of the world at diverse scales. There are infinitely many different ways of abstracting such models, so the particular set of concepts derived needs to relate to a particular map purpose. In this sense, map generalisation can be viewed as facilitating an exchange of knowledge between the map producer and the map user by constructing a particular view of the world (Mustière *et al*, 1999). To achieve this, ontologies (Fonseca *et al,* 2002) can provide a mechanism by which the concepts inherent in this knowledge can be explicitly represented. Guarino (1998) describes four types of ontologies: top level, domain, task, and user. For map generalisation;

- Top-level ontologies could include; general concepts of space and spatial relations, gestalt principles for perceptual grouping (DeLucia and Black, 1987) or the Micro/Meso/Macro agent decomposition (Ruas, 1999).
- Domain-level ontologies might include; network structures (Heinzle *et al*, 2005), urban forms (Lynch 1960, Boffet, 2000),
- Task-level ontologies might describe the goals of different users groups in relation to particular tasks such as recreation or navigation (Kulik, 2005)
- Application ontologies would draw on domain and task ontologies to create models for particular applications. For example, a small scale road map for drivers or a 'point-of-interest' map for tourists.

The main reason for using an ontology is to define the abstract geographic entities that will be present in the generalised map and their inter-relationships and logical entailments. This allows generalisation operations to be performed and constraints described using a single logical model. Users access this model by matching data in their own local schemas to entities in the application ontology. Figure 5 shows a simplified model of an ontology to illustrate how concepts for generalised urban topography might be organised. Two separate domain ontologies are used; one for roads (shown in blue) and one for buildings (shown in red). New concepts are created from the integration of these (shown in green).

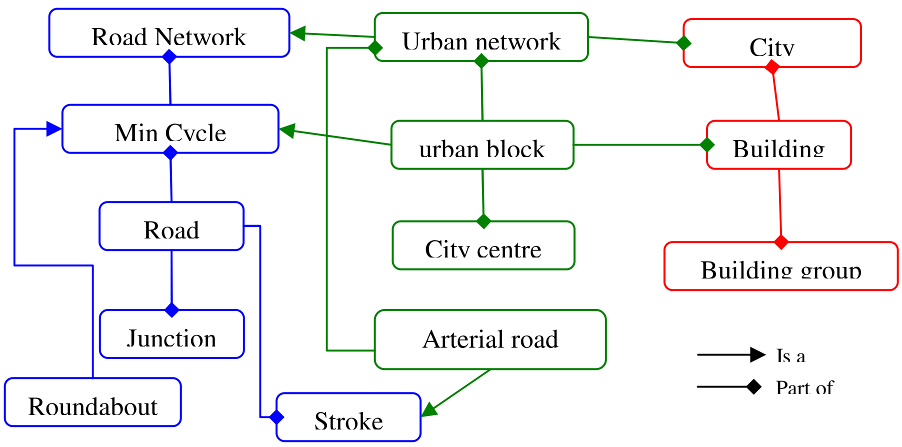

Figure 5 Simplified model of an ontology for urban generalisation

In order to be useful the classes in Figure 5 needs to be instantiated with real world objects having geographic footprints. This allows more complex relationships to be defined amongst classes allowing more expressive semantic constraints. For example, "arterial roads *connecting* city centre should be preserved" or "buildings *in* the city centre should be represented as blocks". Creating entities is part of the matching process between a user's own data schema and the application ontology. Commonly some objects will need to be derived from aggregations of other objects during this process. In the example given in Figure 5 it is possible to geometrically create every other the spatial representations starting only with road and buildings objects. Table 4 illustrates how this has been achieved by different researchers.

Abstract Class	Operation	References
Road Network	Union roads at end points (i.e build topology)	Mackaness, 1995
Min Cycles	Union cycles of roads (e.g. create partitions)	Ruas, 1999
Strokes	Union visually continuous roads segments	Thompson and Richardson, 1999
Junctions	Analysis of node configurations	Mackaness and Mackechnie, 1999
Roundabout	Subset min cycles by area and compactness	
City	Buffer and union buildings	Boffet, 2000
Building group	Analyse and subset buildings (e.g. according to function or alignment)	Christophe and Ruas, 2002 Ruas and Holzapfel, 2003 Burghardt and Steiniger, 2005
Urban Network	Intersect road network and cities	Edwardes and Mackaness, forthcoming
Urban Block	Intersect min cycles in cities with buildings	Ruas and Plazanet 1996
City centre	Analyse and subset urban blocks	Edwardes and Mackaness, forthcoming
Arterial Roads	Intersect strokes with urban network	Heinzle *et al*, 2005

Table 4 Instantiation of entities

Figure 6 illustrates some of these operations for a small city. Figure 7 illustrates the definition of a city center based on statistics for minimum cycle sizes and shows how this entity is preserved by differentially generalizing it.

a) b) c)

Figure 6 a) City and road network, b) urban network c) arterial roads (Edwardes and Mackaness, forthcoming)

a) B

Figure 7 a) Min cycles and city center b) Generalised urban network using the city centre to balance density contrasts (Edwardes and Mackaness, forthcoming)

4 Conclusions

There is a clear motivation within the map generalisation research for systems that can allow the sharing of work. Significant progress has been made in achieving this goal physically by using the web services model. The main barriers to interoperability are now in defining generalisation at the service interface. Progress has been made by integrating standards for encoding geographic representation such as GML within the web services framework. However formalisation of specialised data types also needs to be performed within the generalisation community. A number of researchers have suggested models to meet these needs so the next development will be to harmonise these to a common model, through a consensus standardisation process. Semantic interoperability is harder to accomplish, the construction of domain ontologies is a possible method for achieving this. Formalisation at this level will likely only be fully achievable as the platform develops and researchers interact, discovering the areas in which they need to focus.

References

BADER, M., BARRAULT, M., REGNAULD, N., MUSTIERE, S., DUCHENE, C.,RUAS, A., FRITSCH, E., LECORDIX, F., AND BARILLOT, X. 1999. AGENT Workpackage D2 - Selection of Basic Algorithms. Technical report, University of Zurich.

BEARD, K. 1991. Constraints on rule formation. In Buttenfield, B. and McMaster, R., editors, *Map Generalization:* Making Rules for Knowledge Representation, pages 121–135. Longman, London.

BOFFET A., 2000 Creating urban information for cartographic generalization International Symposium on Spatial Data Handling (SDH), pp 3b4-3b16 , 2000

BURGHARDT, D. and S. STEINIGER; 2005: Usage of Principal Component Analysis in the Process of Automated Generalisation. Proceedings of 22nd International Cartographic Conference, La Coruña, Spain.

BURGHARDT, D.; NEUN, M. and R. WEIBEL; 2005: Generalization Services on the Web – A Classification and an Initial Prototype Implementation. Proceedings American Congress on Surveying and Mapping, Auto-Carto, Las Vegas, USA.

CHRISTOPHE, S. and A. RUAS; 2002: Detecting building alignments for generalisation purposes. In: D. Richardson and P. van Oosterom (eds.), Advances in Spatial Data Handling. 10th Int. Symp. on Spatial Data Handling, Berlin Heidelberg: Springer Verlag, pp. 419–432.

DELUCIA, A. A. AND BLACK, R. T. 1987. A comprehensice approach to automatic feature generalization. *Proceedings of the International Cartographic Conference, Morelia, Mexico*, pages 169–192.

EDWARDES, A.; BURGHARDT, D.; BOBZIEN, M.; HARRIE L.; LEHTO L.; REICHENBACHER, T.; SESTER, M. and R. WEIBEL; 2003: Map Generalisation Technology: Addressing the need for a common research platform. Proceedings of 21st International Cartographic Conference. Durban, South Africa.

EDWARDES, A AND MACKANESS, ForthComing, Road network simplification in urban areas, Computers, Environment and Urban Systems

GML, 2004, OpenGIS® Geography Markup Language (GML) Implementation Specification http://www.opengis.org/techno/implementation.htm (2004 accessed)

GUARINO N. 1998 Formal Ontology and Information Systems. in: N. Guarino (Ed.), *Formal Ontology in Information Systems*. pp. 3-15, IOS Press: Amsterdam, Netherlands.

HARDY, P., HAYLES, M. and REVELL, P., 2003. Clarity - A New Environment for Generalisation using Agents, Java, XML and Topology. ICA Generalisation Workshop, Paris.

HAKE, G.; GRÜNREICH, D. and MENG, L; 2002: Kartographie, deGruyter.

HEINZLE, F.; SESTER, M. and ANDERS, K.; 2005: Graph based approach for recognition of patterns and implicit information in road networks. Proceedings of 22nd International Cartographic Conference, La Coruña, Spain.

JUMP 2005 The Jump project http://www.jump-project.org (accessed july 2005)

KULIK, DUCKHAM, M and EGENHOFER.M 2005. Ontology-driven map generalization. Journal of Visual Languages and Computing, in press, vol. 16, no. 2.

LAMY, S., RUAS, A., DEMAZEAU, Y., JACKSON, M., MACKANESS, W., and WEIBEL, R. 1999. The Application of Agents in Automated Map Generalisation In: Proceedings of the 19th ICA/ACI Conference, pp. 1225{1234. Ottawa.

LYNCH K. 1960 *The Image of the City* (Cambridge: M.I.T. Press, 1960).

MACKANESS W. A., 1995, Analysis of Urban Road Networks to Support Cartographic Generalization. *Cartography and Geographic Information Systems*, **22**, 306 – 316.

MACKANESS W. A. AND MACKECHNIE G. A., 1999, Automating the Detection and Simplification of Junctions in Road Networks. *GeoInformatica*, **3**, 185- 200.

McMASTER, R. B. and SHEA, K. S. (1992). Generalization in Digital Cartography. Association of American Geographers, Wahington D.C.

MUSTIÈRE S, ZUCKER J-D, and SAITTA L, 1999, Cartographic generalization as a combination of representing and abstracting knowledge Proceedings of the 7th ACM international symposium on Advances in geographic information systems

NEUN, M. and S. STEINIGER; 2005: Modelling Cartographic Relations for Categorical Maps. Proceedings of 22nd International Cartographic Conference, La Coruña, Spain.

NEUN, M.; R. WEIBEL and D. BURGHARDT; 2004: Data enrichment for adaptive generalisation. The 7th ICA Workshop on Generalisation and Multiple Representation, Leicester.

PETZOLD, I.; BURGHARDT, D. and M. BOBZIEN; 2005: Automated derivation of town plans from large scale data – on an example of area to line simplification. Proceedings of 22nd International Cartographic Conference, La Coruña, Spain.

REGNAULD, N. 2005 spatial structures to support automatic generalisation, Proceedings of 22nd International Cartographic Conference, La Coruña, Spain.

RUAS, A. 1999. Modele de generalisation de donnees geographiques a base de contraintes et d'autonomie. Ph.D. thesis, Universite de Marne la Vallee.

RUAS, A., PLAZANET, C., 1996., Strategies for automated generalisation, In Proc 7th Int. Symp. on Spatial Data Handling. pp 6.1-6.17.

RUAS, A. and F. HOLZAPFEL; 2003: Automatic characterisation of building alignments by means of expert knowledge. Proceedings of 21st International Cartographic Conference, Durban, South Africa., 2003, pp. 1604-1615.

SESTER, M. 2000. Generalization Based on Least-squares Adjustment. In: Int. Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B4, Amsterdam, 931-938.

SHEPERD, I.D.H. 1991 Information integration and GIS. In David Maguire, Michael Goodchild, and David Rhind (Eds.) Geographical Information Systems - Volume 1: Principles. Longman Scientific and Technical: London.

SOAP, 2003, SOAP Version 1.2 Part 1: Messaging Framework. http://www.w3.org/TR/soap (accessed 01/2005)

THOMSON, R. C., and RICHARDSON, D. E., 1999, The 'Good Continuation' Principle of Perceptual Organization Applied to the Generalization of Road Networks. *Proceedings of the ICA 19th International Cartographic Conference* (Ottawa, 14-21 August 1999), pp.1215–1223.

UDDI, 2004, Universal Description, Discovery, and Integration. http://www.uddi.org/ (accessed 01/2005)

VCKOVSKI, A., 1998, Interoperable and Distributed Processing in GIS. (London:Taylor & Francis)

WARE, J.M., C.B. JONES and N. THOMAS. 2003. Automated Map Generalization with Multiple Operators: A Simulated Annealing Approach. International Journal of Geographical Information Science, 17(8): 743-769.

WSA, 2004, Web Services Architecture. http://www.w3.org/TR/ws-arch/ (accessed 01/2005)

WSDL 2001, Web Services Description Language (WSDL) 1.1. http://www.w3.org/TR/wsdl (accessed 01/2005)