Collaboration on an Ontology for Generalisation

Gould, N.M.[1] Mackaness, W.A.[2] Touya, G.[3] Hart, G.[4]

[1]Geography and Environmental Management, Manchester Metropolitan University
Email: n.gould@mmu.ac.uk

[2]School of GeoSciences, The University of Edinburgh
Email: william.mackaness@ed.ac.uk

[3]COGIT Laboratory, IGN
Email: Guillaume.Touya@ign.fr

[4]University of Nottingham
Email: Glen.Hart@nottingham.ac.uk

To move beyond the current plateau in automated cartography we need greater sophistication in the process of selecting generalisation algorithms. This is particularly so in the context of machine comprehension. We also need to build on existing algorithm development instead of duplication. More broadly we need to model the geographical context that drives the selection, sequencing and degree of application of generalisation algorithms. We argue that a collaborative effort is required to create and share an ontology for cartographic generalisation focused on supporting the algorithm selection process. The benefits of developing a collective ontology will be the increased sharing of algorithms and support for on-demand mapping and generalisation web services.

1 Introduction

An ontology can be defined as a “formal, explicit specification of a shared conceptualization” (Studer et al., 1998). The term “shared”, first introduced by Borst (1997), implies that the knowledge encapsulated is consensual, without which the benefits of the ontology are limited. This is why the development of any ontology has to be a collaborative effort. Here we call for a collective effort to build an ontology for generalisation.

The proposal is to extend and broaden the work of Gould and Cheng (2013) where an ontology was designed specifically to support the on-demand mapping of road accidents. The aim is to collaboratively develop an ontology for the selection of generalisation algorithms.

Consider the number of algorithms that perform road network generalisation (Benz and Weibel, 2013; Weiss and Weibel, 2013; ESRI, 2012; Li and Zhou, 2012; Yang et al., 2011; Liu et al., 2010; Savino et al., 2010; Touya, 2010; Chen et al., 2009; Jiang and Claramunt, 2004; Thomson and Richardson, 1999; Mackaness, 1995). It requires a reading of the literature to determine which of these is appropriate for a particular context. But if it is difficult for humans to select an algorithm based on a natural language description then it is even more so for machines. If the move to develop spatial data infrastructures based on web services continues then there is a need to formalise algorithm knowledge to aid their automatic selection.

2 Why develop an ontology of generalisation?

Although the results of generalisation are the results of manipulating geometric primitives, those manipulations need to be based on the geographical context – not just user needs (Mackaness, 2007). When mapping road accidents, for example, the road network cannot be generalised without taking account of its relationship with the location and type of accidents. For example, the generalisation of road network might lead to the removal of a minor road that provides context for a group of mapped accidents; road accidents mapped without the road on which they occur are meaningless. Context is part of the *semantics* of a domain, which can be "encapsulated, elucidated, and specified by an ontology" (Kavouras and Kokla, 2008, p10). In the domain of generalisation, the semantic relationships that govern generalisation must be made explicit and formalised (Wolf, 2009; Dutton and Edwardes, 2006). In this example, the ontology would describe the semantic relationship between road accidents and the road network. Over twenty years ago, (Nyerges, 1991) pointed out that cartographers lacked the means to systematically document the knowledge required for generalisation. The concept of ontologies provides that means.

Using an ontology to describe a domain can lead to intelligent knowledge retrieval (Benjamins et al., 1998). In particular, if the ontology is formalised using logic, such as the Web Ontology Language (OWL), automated reasoning is possible (Horrocks, 2013). The aim is to develop an ontology that sufficiently describes the characteristics of generalisation algorithms, the geographic features they operate on, and the context they operate in, such that algorithms can be selected automatically. Such characteristics might include the effects the algorithm has on the geographic features, such as a change in dimensionality or a reduction in the number of features, and source and target scales. To express these characteristics, entities such as *Operator*, *Algorithm*, *Geometry*, *Feature Type* and *Feature Collection* will need to be encapsulated in the ontology. These entities will then be related using axioms such as *Collapse reduces Dimensionality*.

It has been argued that every information system contains an implicit ontology and making it explicit reduces conflicts between the ontological concepts and the implementation (Fonseca et al., 2002). For example, two different algorithm programmers might have different understandings of the concept of the *amalgamation* and thus create different implementations that share the same name. The ontology can be used to describe characteristics of each algorithm, the input features and the current context. If the descriptions contain sufficient detail then the appropriate algorithm can be selected using its characteristics and not its name.

The knowledge required for automatic generalisation is currently encapsulated either *formally* and *implicitly* (Figure 1) in generalisation software such as agent-based systems (Taillandier and Taillandier, 2012) or *informally* and *explicitly*, for example, in different generalisation operation taxonomies (Roth et al., 2011; Foerster et al., 2007; McMaster and Shea, 1992). Encapsulating that knowledge explicitly and formally in an ontology will allow that knowledge to be shared, expanded and utilised by generalisation software for all forms of map production and interaction. The formalisation of the description of the features we wish to map will also lead to "smarter" data where the business logic is no longer held in the software but with the data (Carral et al., 2013); how the features are described in the ontology will determined how they are generalised.

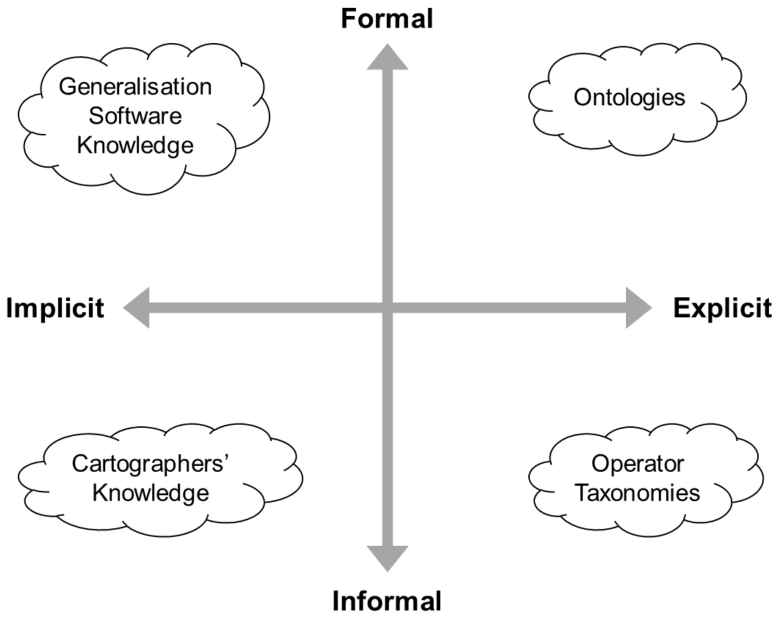

Figure 1 Classifying generalisation knowledge representations

Ontologies have also been used in generalisation by Kulik et al. (2005) to aid road line simplification; by Dutton and Edwardes (2006) to represent the roles of geographic features and semantic and structural relationships between features in a coastal region; by Wolf (2009) to influence the aggregation and dimensional collapse of features; and by Lüscher et al. (2008) to aid the recognition of terrace houses. However, none of these ontologies described the *process* of generalisation; the identification, sequencing and execution of generalisation algorithms. A collaborative effort is called for in the creation of such an ontology.

For an ontology to be successful, if success is measured by its acceptance and use within its target community, then the creation of the ontology has to be a collaborative process. Indeed the development of an ontology - a *shared* conceptualisation - is *intrinsically* a collaborative process (Groza et al., 2013). A collaboration will have a plurality of stakeholders and is therefore more likely to lead to a sustainable ontology. However, in a collaboration involving partners with potentially differing perspectives and goals, it is important to define a clear methodology for the ontology development.

3 Methodology

The application of a methodology will provide structure to the process of building an ontology and will ensure best practice (Hart and Dolbear, 2013). A number of ontology design methodologies, which provide guidance on defining concepts and relationships, have been developed but there is a lack of widely accepted methodologies, possibly because many of them were developed for a particular application (Iqbal et al., 2013).

However, a number of criteria to aid selection of an ontology methodology have been defined (Iqbal et al., 2013) and one important criterion is the support provided by the methodology for *collaborative construction*, where different members of the team, even in different geographic locations, can work simultaneously on different aspects of the ontology.

The much cited Ontology Development 101 guide of Noy and McGuinness (2001)is a methodology in all but name. Their guide states that for an ontology to work effectively it should be designed with a specific task in mind. It is necessary therefore to be clear about the purpose of the ontology before the process of designing it is started. The aim of the proposed project is not the design of an all-purpose ontology but an ontology to aid the selection of generalisation algorithms.

Many methodologies stress the importance of using existing ontologies and there has been recent interest in "geo-ontologies" (Janowicz et al., 2012), which might be of use, particularly when describing physical geographic concepts. However, the focus will be on those characteristics of geographic features that effect their generalisation. Immaterial concepts such as generalisation *operator* and *algorithm* will also need to be defined and it may be that a set of related ontologies, developed by different partners but to a common template, will need to be developed rather than a single ontology.

The open source webProtégé ontology editor (Tudorache et al., 2013) allows for the collaborative development of ontologies and can be downloaded and installed on a server or used via the Stanford University hosted version. Collaboration is supported by change tracking, and a discussion forum for each class defined in the ontology (Figure 2). Ontologies developed using webProtégé support the popular OWL 2 standard for ontologies (Grau et al., 2008), which will be employed by the project.

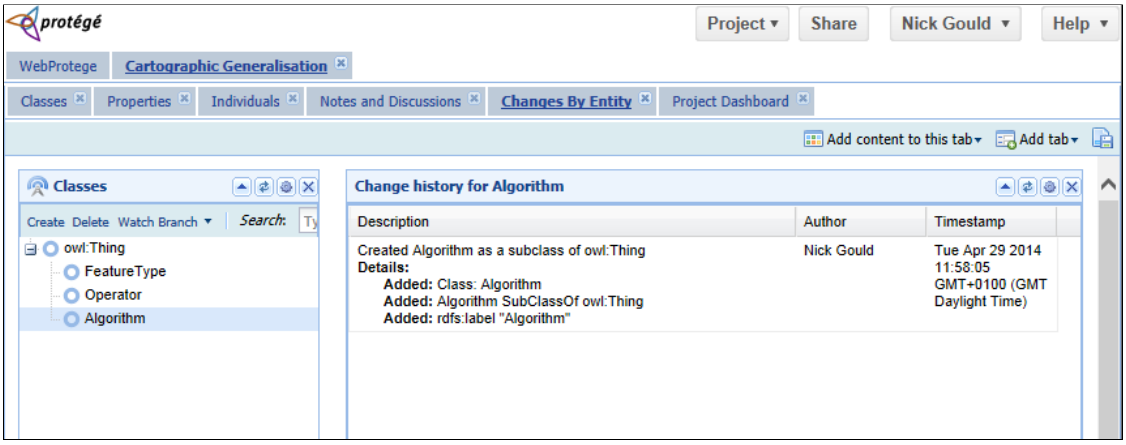

Figure 2 The webProtégé interface

The webProtégé interface is specifically a tool for the knowledge engineer. It is important to recognise the distinctly separate roles of the knowledge engineer and the domain expert and the key collaboration is between these two roles. It might be possible for domain experts to act as knowledge engineers (Denaux et al., 2011).

4 Evaluating and using the ontology

The evaluation of the ontology is seen as a key phase in several ontology development methodologies (Sure et al., 2009; Grüninger et al., 2008; Fernández-López et al., 1997). We propose a number of measures, of increasing complexity, to evaluate the ontology. Firstly, a set of *competency questions* (Grüninger and Fox, 1995), which are proposed at the start of the design process, can be used to evaluate whether the ontology meets requirements and act as a justification for the ontology. In particular, the competency questions will help define the scope of the ontology.

Once the competency questions have been answered to satisfaction the second stage in the evaluation will be to use the ontology to capture the properties of a number of generalisation algorithms described in the literature. Particular types of algorithm such as road network generalisation or building amalgamation can be used as case studies. Line simplification algorithms would provide a particularly interesting use case since there are a large number of competing algorithms, which often share very similar, but distinctly different, characteristics. The ontology is evaluated by comparing the algorithm definitions; if any two algorithms have exactly the same definition then their definitions need to be refined, possibly by introducing new concepts and relationships to the ontology. It is also necessary to consider the fact that some algorithms may implement multiple operators such as aggregation *and* simplification. The next two evaluation steps will put the ontology to practical use, firstly by using it to support web-based generalisation services.

For geospatial web services to be used effectively they require *semantic interoperability*: "the ability of services to exchange data in a meaningful way and with a minimum of human intervention" (Janowicz et al., 2010, p112). The standard for implementing geospatial web services is the OGC's Web Processing Service (WPS) protocol, which has been adopted by the WEBGEN generalisation platform (Neun et al., 2013). The WPS protocol allows for a natural language description of what the service provides but not the formal, machine readable description required for semantic interoperability. Semantic interoperability allows services to interact at a semantic level, not just a syntactic level.

The WPS protocol allows for syntactic interoperability, by requiring the formal definition of the input and output parameters (specifying geometries, for example) but again the semantics are missing; why should one building displacement algorithm be chosen over another? The knowledge required to make such decisions can be formalised using an ontology, but how can the ontology be integrated with generalisation web services?

One solution is to introduce a *Semantic Enablement Layer* (Janowicz et al., 2010), where a *Web Ontology Service* injects semantics into both data and processing services (Figure 3). The Web Ontology Service could be employed to maintain a shared set of generalisation ontologies. In this depiction a separate ontology is used to describe geographic features and it may be that a set of related ontologies will need to be developed rather than a single ontology. There might, for example, be a separate ontology for describing spatial relations (Touya et al., 2014). A *Web Reasoning Service* can then be used to search for an appropriate generalisation service by using Web Ontology Language (OWL) queries. We suggest the application of a Semantic Enablement Layer to the WEBGEN service (Dresden University of Technology, 2014).

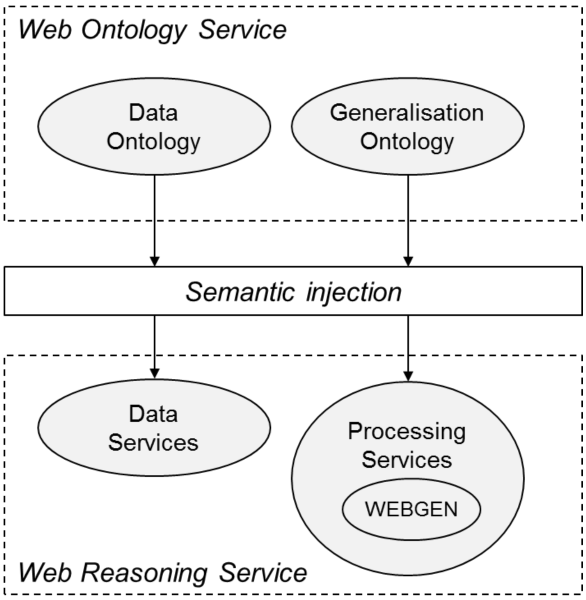

Figure 3 Semantic injection of data together with processing web services

The final evaluation of the ontology will be to utilise it to support existing map generalisation software such as an agent-based system. As stated earlier, much domain knowledge is embedded in generalisation software systems and the formalisation of some of that knowledge in an ontology can help make it shareable. Currently such systems are configured to generalise familiar feature types such as roads, rivers and buildings at a fixed set of target scales and their associated knowledge base has to be updated every time a new generalisation algorithm is introduced to the system or when user requirements change (Taillandier and Taillandier, 2012). If the knowledge required to generalise road accidents, for example, is shared via an ontology, then such systems might become more flexible.

5 Conclusion

The aim of the project is to develop an ontology of cartographic generalisation. Specifically, the ontology should contain sufficient knowledge to allow for the automatic selection of generalisation algorithms. To achieve this, we propose the creation of a consortium of partners, and collective agreement on a project plan to confirm the aims of the project, define a methodology for developing the ontology, and define criteria for evaluating the ontology. The ontology should build on previous attempts at formalising generalisation knowledge such as for map specifications and constraints (Stoter et al., 2010; Burghardt et al., 2007).

In this paper we have argued the need for a generalisation ontology for algorithm selection of a form that supports machine reasoning, and thus higher levels of automation in the delivery of web-based generalisation services. Such an effort would reduce redundancy in the development of algorithms. More broadly this ambition reflects the need for ‘geographic generalisation’ methodologies – i.e. algorithms sensitive to the geographic context of the problem rather than just the geometric.

References

BENJAMINS, V. R., FENSEL, D. & GOMEZ PEREZ, A. 1998 Knowledge Management through Ontologies. *In Second International Conference on Practical Aspects Knowledge Management*. Basel, Switzerland, 29th-30th October 1998.

BENZ, S. A. & WEIBEL, R. 2013 Road network selection using an extended stroke-mesh combination algorithm. *In 16th Workshop of the ICA commission on Generalisation and Multiple Representation*. Dresden, 23rd–24th August 2013.

BORST, W. 1997. *Construction of Engineering Ontologies.* PhD, University of Twente.

BURGHARDT, D., SCHMID, S. & STOTER, J. 2007 Investigations on Cartographic Constraint Formalisation. *In 10th Workshop of the ICA commission on Generalisation and Multiple Representation*. Moscow, Russia, 2nd-3rd August 2007.

CARRAL, D., SCHEIDER, S., JANOWICZ, K., VARDEMAN, C., KRISNADHI, A. & HITZLER, P. 2013. An Ontology Design Pattern for Cartographic Map Scaling. *In:* CIMIANO, P., CORCHO, O., PRESUTTI, V., HOLLINK, L. & RUDOLPH, S. (eds.) *The Semantic Web: Semantics and Big Data.* Springer: Berlin/Heidelberg.

CHEN, J., HU, Y., LI, Z., ZHAO, R. & MENG, L. 2009. Selective omission of road features based on mesh density for automatic map generalization. *International Journal of Geographical Information Science,* 23(8) pp.1013-1032.

DENAUX, R., DOLBEAR, C., HART, G., DIMITROVA, V. & COHN, A. G. 2011. Supporting domain experts to construct conceptual ontologies: A holistic approach. *Web Semantics: Science, Services and Agents on the World Wide Web,* 9(2) pp.113-127.

DRESDEN UNIVERSITY OF TECHNOLOGY. 2014. *WebGen-WPS Portal* [Online]. Available: http://kartographie.geo.tu-dresden.de/webgen_docs/ [Accessed 12th June 2014].

DUTTON, G. & EDWARDES, A. 2006 Ontological Modeling of Geographical Relationships for Map Generalization. *In Workshop of the ICA Commission on Map Generalization and Multiple Representation*. Portland, USA, 25th June 2006.

ESRI. 2012. *Thin Road Network (Cartography)* [Online]. Available: http://resources.arcgis.com/en/help/main/10.1/index.html#/Thin_Road_Network/007000000014000000/ [Accessed 12th June 2014].

FERNÁNDEZ-LÓPEZ, M., GÓMEZ-PÉREZ, A. & JURISTO, N. Methontology: from ontological art towards ontological engineering. Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series, 24-26 March 1997 1997 Stanford University. American Asociation for Artificial Intelligence.

FOERSTER, T., STOTER, J. & KÖBBEN, B. 2007 Towards a formal classification of generalization operators. *In International Cartographic Conference 2007*. Moscow, 4-10 August, 2007.

FONSECA, F., EGENHOFER, M., DAVIS, C. & CÂMARA, G. 2002. Semantic Granularity in Ontology-Driven Geographic Information Systems. *Annals of Mathematics and Artificial Intelligence*, 36(1-2) pp.121-151.

GOULD, N. & CHENG, J. 2013 A prototype for ontology driven on-demand mapping of urban traffic accidents. *In 16th Workshop of the ICA commission on Generalisation and Multiple Representation*. Dresden, 23rd–24th August 2013.

GRAU, B. C., HORROCKS, I., MOTIK, B., PARSIA, B., PATEL-SCHNEIDER, P. & SATTLER, U. 2008. OWL 2: The next step for OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4) pp.309-322.

GROZA, T., TUDORACHE, T. & DUMONTIER, M. 2013. State of the art and open challenges in community-driven knowledge curation. *Journal of Biomedical Informatics*, 46(1) pp.1-4.

GRÜNINGER, M., BODENREIDER, O., OLKEN, F., OBRST, L. & YIM, P. 2008. Ontology Summit 2007 – Ontology, taxonomy, folksonomy: Understanding the distinctions. *Applied Ontology*, 3(3) pp.191-200.

GRÜNINGER, M. & FOX, M. 1995 Methodology for the Design and Evaluation of Ontologies. *In IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*. Montreal, 19-20 August 1995.

HART, G. & DOLBEAR, C. 2013. *Linked Data: A Geographic Perspective,* Boca Raton, USA, CRC Press.

HORROCKS, I. 2013. What Are Ontologies Good For? *In:* KÜPPERS, B.-O., HAHN, U. & ARTMANN, S. (eds.) *Evolution of Semantic Systems.* Berlin/Heidelberg: Springer.

IQBAL, R., MURAD, M. A. A., MUSTAPHA, A. & SHAREF, N. M. 2013. An Analysis of Ontology Engineering Methodologies: A Literature Review. *Research Journal of Applied Sciences, Engineering and Technology*, 16(6).

JANOWICZ, K., SCHADE, S., BRORING, A., KESSLER, C., MAUÉ, P. & STASCH, C. 2010. Semantic Enablement for Spatial Data Infrastructures. *Transactions in GIS,* 14(2) pp.111.

JANOWICZ, K., SCHEIDER, S., PEHLE, T. & HART, G. 2012. Geospatial semantics and linked spatiotemporal data – Past, present, and future. *Semantic Web,* 3(4) pp.321-332.

JIANG, B. & CLARAMUNT, C. 2004. A Structural Approach to the Model Generalization of an Urban Street Network. *Geoinformatica,* 8(2) pp.157-171.

KAVOURAS, M. & KOKLA, M. 2008. *Theories of geographic concepts : ontological approaches to semantic integration*, Boca Raton, CRC Press.

KULIK, L., DUCKHAM, M. & EGENHOFER, M. 2005. Ontology-driven map generalization. *Journal of Visual Languages & Computing,* 16(3) pp.245-267.

LI, Z. & ZHOU, Q. 2012. Integration of linear and areal hierarchies for continuous multi-scale representation of road networks. *International Journal of Geographical Information Science*, 26(5) pp.855-880.

LIU, X., ZHAN, F. B. & AI, T. 2010. Road selection based on Voronoi diagrams and strokes in map generalization. *International Journal of Applied Earth Observation and Geoinformation*, 12(Supplement 2) pp.S194-S202.

LÜSCHER, P., WEIBEL, R. & MACKANESS, W. A. 2008. Where is the Terraced House? On the Use of Ontologies for Recognition of Urban Concepts in Cartographic Databases. *In:* RUAS, A. & GOLD, C. (eds.) *Headway in Spatial Data Handling.* Berlin/Heidelberg: Springer.

MACKANESS, W. 1995. Analysis of Urban Road Networks to Support Cartographic Generalization. *Cartography and Geographic Information Systems,* 22(4) pp.306-316.

MACKANESS, W. A. 2007. Understanding Geographic Space. *In:* MACKANESS, W. A., RUAS, A. & SARJAKOSKI, L. T. (eds.) *Generalisation of Geographic Information.* Amsterdam: Elsevier Science B.V.

MCMASTER, R. B. & SHEA, K. S. 1992. *Generalization in digital cartography,* Washington D.C., Association of American Geographers.

NEUN, M., REGNAULD, N. & WEIBEL, R. 2013. WEBGEN: Web Services to Share Cartographic Generalization Tools. *Innovative Software Development in GIS.* John Wiley & Sons, Inc.

NOY, N. F. & MCGUINNESS, D. 2001. Ontology Development 101: A Guide to Creating Your First Ontology. Technical Report SMI-2001-0880, Stanford University.

NYERGES, T. L. 1991. Representing geographical meaning. *In:* BUTTENFIELD, B. & MCMASTER, R. B. (eds.) *Map Generalization: Making rules for knowledge representation.* New York: Longman Scientific & Technical.

ROTH, R., BREWER, C. & STRYKER, M. 2011. A typology of operators for maintaining legible map designs at multiple scales. *Cartographic Perspectives*, 68.

SAVINO, S., RUMOR, M., ZANON, M. & LISSANDRON, I. 2010 Data enrichment for road generalization through analysis of morphology in the cargen project. *In 13th Workshop of the ICA commission on Generalisation and Multiple Representation.* Vol. 12th-13th September 2010. Zurich, Switzerland.

STOTER, J., BAELLA, B., BLOK, C., BURGHARDT, D., DUCHÊNE, C., PLA, M., REGNAULD, N. & TOUYA, G. 2010. *State-of-the-art of automated generalisation in commercial software*, Dublin, EuroSDR.

STUDER, R., BENJAMINS, R. & FENSEL, D. 1998. Knowledge engineering: Principles and methods. *Data and Knowledge Engineering,* 25(1-2) pp.168-198.

SURE, Y., STAAB, S. & STUDER, R. 2009. Ontology Engineering Methodology. *In:* STAAB, S. & STUDER, R. (eds.) *Handbook on Ontologies.* 2nd ed. Berlin/Heidelburg: Springer-Verlag.

TAILLANDIER, P. & TAILLANDIER, F. 2012. Multi-criteria diagnosis of control knowledge for cartographic generalisation. *European Journal of Operational Research,* 217(3) pp.633-642.

THOMSON, R. & RICHARDSON, D. 1999 The ‘Good Continuation’ Principle of Perceptual Organization applied to the Generalization of Road Networks. *In 17th International Cartographic Conference*. Ottawa.

TOUYA, G. 2010. A Road Network Selection Process Based on Data Enrichment and Structure Detection. *Transactions in GIS*, 14(5) pp.595-614.

TOUYA, G., BUCHER, B., FALQUET, G., JAARA, K. & STEINIGER, S. 2014. Modelling Geographic Relationships in Automated Environments. *In:* BURGHARDT, D., DUCHENE, C. & MACKANESS, W. A. (eds.) *Abstracting Geographic Information in a Data Rich World.* Heidelberg: Springer.

TUDORACHE, T., NYULAS, C., NOY, N. F. & MUSEN, M. A. 2013. WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the Web. *Semantic Web,* 4(1) pp.89-99.

WEISS, R. & WEIBEL, R. 2013 Road Network Selection for Small-Scale Maps Using an Improved Centrality Approach. *In 16th Workshop of the ICA commission on Generalisation and Multiple Representation*. Dresden, 23rd–24th August 2013.

WOLF, E. 2009. Ontology-Driven Generalization of Cartographic Representations by Aggregation and Dimensional Collapse. *In:* BERNSTEIN, A., KARGER, D., HEATH, T., FEIGENBAUM, L., MAYNARD, D., MOTTA, E. & THIRUNARAYAN, K. (eds.) *The Semantic Web - ISWC 2009.* Springer Berlin / Heidelberg.

YANG, B., LUAN, X. & LI, Q. 2011. Generating hierarchical strokes from urban street networks based on spatial pattern recognition. *International Journal of Geographical Information Science*, 25(12) pp.2025-2050.