Questionnaires for Ontology Engineering

Marie-Kristina Thomson

Department of Geomatic Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom
thomson@ge.ucl.ac.uk

Abstract. Questionnaires provide a useful instrument for capturing knowledge as part of ontology engineering. Often knowledge is derived from a single domain expert who may be personally involved in the ontology project. However, in domains of common phenomena such as geography knowledge can be drawn from larger populations. This paper presents a number of implications for such an endeavour based on the author's personal experiences that highlight aspects of the nature of knowledge, categorisation and conceptualisation, and cultural and linguistic constraints.

1 INTRODUCTION

Social sciences apply both qualitative and quantitative research methods in order to explore human behaviour (Pope and Mays, 1995). In ontology engineering the concern lies with formalising a conceptualisation of a specific domain for which recording of relevant knowledge is necessary. Finding a source of expertise that can be harvested is one big challenge, the other is faced by how this information can be most efficiently extracted. Indeed, most knowledge acquisition techniques focus on interviewing a single domain expert, or extracting knowledge from loosely structured textual or multimedial data, or databases (Svátek, 2006). However, if such knowledge is to be drawn from a wider population, then survey techniques as employed by social sciences become a means to an end (Richie and Lewis, 2003; Patton, 1990; Green and Thorogood, 2004).

This paper was initiated through personal experiences from gathering knowledge for the land use domain (Thomson and Béra, 2007). In section 2 the nature of knowledge is discussed with respect to the expert in a geographical context. Section 3 presents relevant issues for capturing knowledge through questionnaires, before conclusions are drawn.

2 WHO IS THE EXPERT?

The relation between humans and their knowledge about and their interaction with the environment is not an easy one to ground, as can be seen from

numerous theories such as empiricism, positivism, rationalism, idealism, or constructivism, that offer different explanations for the nature of knowledge (Gale and Golledge, 1982). Apart from technical knowledge that may be well documented, what is individually understood as knowledge depends largely on a person's beliefs and truths that one confides in. Thus, what is accepted as knowledge may well be infiltrated by false beliefs and truths (Smith, 2004). Especially if a system of knowledge relates to the geographical domain, subjects become malleable to physical, cultural and social influences present in that environment. As beliefs may vary according to gender, ethnic, cultural differences, and spatial context, it is insufficient to transform beliefs directly into a domain knowledge.

Nevertheless, as Wilson and Keil (1999) argue, building a knowledge-based system is like intellectual cloning, where one finds persons with a reasoning skill that is important to the problem at hand, talk to them to determine what specialised knowledge they have and how they reason, then embody that knowledge and reasoning in a program. Usually the choice of expert is determined by their knowledgeability in the domain of concern. In the geographical realm, where the environment poses as the normal setting for people's activities (Downs and Stea, 1977) and knowledge is concerned with common, natural phenomena, every person becomes an expert in their own rights (Lowenthal, 1961). This allows to gather manifold, context-rich observations which not only aid the grounding of ontologies (Kuhn, 2001) but ensure validity by generalising subjective knowledge to the objective reality (Thomson and Béra, 2007).

3 CAPTURING KNOWLEDGE

The usual guidelines apply for developing questionnaires (Oppenheim, 1992; Willis, 2005; Taylor-Powell, 1999), that hold commonly known difficulties in administering questionnaire surveys (Boynton and Greenhalgh, 2004; Wall *et al*., 2002). According to Svátek (2006) questionnaires provide a useful resource in ontology design due to their small extent and richness in knowledge, and their relevance in terms of the domain and its applications. However, the greatest challenge lies with uncovering one's knowledge, in particular since people may be unaware of their knowledge (Downs and Stea, 1977), and have no direct access to their own unconscious processes. For example, in Thomson and Béra (2007) different methods were used to access knowledge by developing a conceptualisation task founded on Rosch's (1978) proposed category system and studying reasoning abilities based on the processes that took place on a topographic map whilst being read and interpreted for land use types by respondents. Although the mind is compelled to organise and structure information to

achieve cognitive economy (Pinker, 1997), the capturing of knowledge that does not directly touch upon respondents' conscious daily habits and beliefs demands great efforts from a respondent's perspective. However, a sample's suitability and skill to perform certain tasks can usually be ensured through purposive sampling strategies (Pope and Mays, 1995).

Another issue is language which provides the means to convey information. The impreciseness and ambiguities of natural languages prohibit their direct transformation into formal languages of artificial construct and defined rules (Tversky, 1993). Furthermore, due to its determining power of how the world is perceived and conceptualised, language is at the root of many different worldviews (Sperber and Hirschfeld, 1999). Despite sceptical views regarding the significance of cultural differences in spatial cognition (e.g. Montello, 1995), human mental processes make use of cultural tools, and are a reflection of that culture just as they are a reflection of the spatial environment at that place and time. Therefore, respondents' roots of nationality and their spoken languages need be accounted for possible differences in beliefs and concepts (Mark and Turk, 2003).

4 CONCLUSION

The aim of ontology engineering is to find a coherent means by which concepts and their characteristics can span the divide between concepts as creatures of the mind and as properties of objects in the world (Smith *et al*., 2005). No matter which domain the ontology addresses, the building of a conceptualisation worthy of representing real world entities according to their interaction with human beings will lead to more human-conceptible models. The danger, however, in any such qualitative approach lies in the infiltration of human knowledge with erroneous beliefs. This can only be erased by generalising subjective views to objective knowledge and accounting for cognitive abilities that can be all the same across cultures, but may also vary depending on domain-specific acquired knowledge.

5 REFERENCES

Boynton, P. M. and Greenhalgh, T. (2004) Hands-on guide to questionnaire research: Selecting, designing, and developing your questionnaire. *BMJ* **328**:1312-1315

Downs, R.M. and Stea, D. (1977) *Maps in Minds - Reflections on Cognitive Mapping*. New York: Harper & Row

Gale, N. and Golledge, R.G. (1982) On the Subjective Partitioning of Space. *Annals of the Association of American Geographers* **72**(1): 60-67

Green, J. and Thorogood, N. (2004) *Qualitative Methods for Health Research.* London: Sage Publications Ltd

Kuhn, W. (2001) Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science* **15**(7): 613-631

Lowenthal, D. (1961) Geography, experience, and imagination: Towards a geographical epistemology. *Annals of the Association of American Geographers* **51**: 241–60.

Mark, D. M. and Turk, A. G. (2003) Landscape categories in Yindjibarndi: Ontology, Environment, and Language. In Kuhn, W., Worboys, M. and Timpf, S. (Eds.): *Spatial Information Theory: Foundations of Geographic Information Science, COSIT'03*. Lecture Notes in Computer Science 2825, Berlin Springer Verlag, pp. 31-49

Montello, D. R. (1995) How significant are cultural differences in spatial cognition? In Frank, A. U. and Kuhn, W. (Eds.): *Spatial Information Theory: A Theoretical Basis for GIS. International Conference COSIT'95,* Semmering, Austria, September 21-23, Proceedings, Lecture Notes in Computer Science 988, Berlin Springer Verlag, pp. 485-500

Oppenheim, A. N. (1992) *Questionnaire Design, Interviewing and Attitude Measurement.* 2nd Ed. London: Continuum

Patton, M. Q. (1990) *Qualitative evaluations and research methods*. (2nd ed.) Newbury Park, CA: Sage

Pinker, S. (1997) *How the mind works*. London: Allen Lane The Penguin Press

Pope, C. and Mays, N. (1995) Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ* **311**:42-45

Ritchie, J. and Lewis, J. (2003) *Qualitative Research Practice*. London: Sage

Rosch, E. (1978) Principles of Categorization. In Rosch, E. and Lloyd, B.B. (Eds.): *Cognition and Categorization*. New Jersey: Lawrence Erlbaum Associates

Smith, B. (2004) Beyond Concepts: Ontology as Reality Representation. In Varzi, A. and Vieu, L. (Eds): *Proceedings of the International Conference on Formal Ontology and Information Systems FOIS 2004*, Turin, 4-6 November

Smith, B., Ceuster, W. and Temmerman, R. (2005) Wüsteria. In Engelbrecht, R., Geissbuhler, A., Lovis, C. and Mihalas, G. (Eds) Connecting

Medical Informatics and Bio-Informatics. *Proceedings of MIE2005 – The XIXth International Congress of the European Federation for Medical Informatics* **116**: 647-652

Sperber, D. and Hirschfeld, L. (1999) *Culture, Cognition, and Evolution*. In Wilson, R.A. and Keil, F.C. (Eds): The MIT Encyclopedia of the Cognitive Sciences, pp. xv-xxxvii, Massachusetts: The MIT Press

Svátek, V. (2006) Ontologies, Questionnaires and (Mining) Tabular Data. In *the 3rd European Semantic Web Conference (ESWC 2006)*, Budva, Montenegro, 11th-14th June

Taylor-Powell, E. (1999) Questionnaire Design: Asking questions with a purpose. G3658-2 Program Development and Evaluation. University of Wisconsin-Extension, Cooperative Extension. Available from URL: http://learningstore.uwex.edu/pdf/G3658-2.pdf [Last Accessed 06/03/2007]

Thomson, M.-K. and Béra, R. (2007) Relating Land Use to the Landscape Character: Toward an Ontological Inference Tool. In Winstanley, A. C. (Ed.): *Proceedings of GIS Research UK Conference, GISRUK'07*. Maynooth, Ireland, 11-13 April 2007, pp. 83-87

Tversky, B. (1993) Cognitive maps, cognitive collages, and spatial mental models. In Frank, A. U. and Campari, I. (Eds): *Spatial Information Theory: A Theoretical Basis for GIS, Proceedings COSIT'93*. Lecture Notes in Computer Science, 716, pp.14-24, Springer: Berlin

Wall, C. R., DeHaven, M. J., and Oeffinger, K. C. (2002) Survey methodology for the uninitiated. *The Journal of Family Practice* **51**(6):1-11

Willis, G. B. (2005) *Cognitive Interviewing. A tool for improving questionnaire design.* California: Sage Publications

Wilson, R. A. and Keil, F. C. (Eds) (1999) *The MIT Encylopedia of the Cognitive Sciences*. Massachusetts: The MIT Press