Preface

Miguel-Ángel Sicilia

Computer Science Department,
University of Alcalá, Ctra. Barcelona, km. 33.6 – 28871,
Alcalá de Henares (Madrid), Spain
E-mail: msicilia@uah.es

Biographical notes: Miguel-Ángel Sicilia obtained his degree in Computer Science from the Pontifical University of Salamanca, Madrid, Spain in 1996 and his PhD from Carlos III University in 1999. From 1997 to 1999 he worked as an Assistant Professor and later as a part-time Lecturer at the Computer Science Department of the same university. He also worked as a Software Architect in e-commerce consulting firms. From 2002 to 2003, he worked as a full-time Lecturer at Carlos III University, after which he joined the University of Alcalá. His research interests are primarily in the areas of adaptive hypermedia, learning technology, and human-computer interaction, with a special focus on the role of uncertainty and imprecision handling techniques in those fields

The diverse views on metadata research

This second issue of IJMSO reflects the diversity of approaches to 'metadata research', ranging from empirical research to the proposal of novel metadata schemas. Such diversity raises questions on scientific methodology that are relevant for metadata and semantics research, and of course, that are critical for IJMSO as a research outlet for a concrete field of study. In what follows, a quick look at some of the methodological aspects present in this second issue of IJMSO is provided. This is intended to provoke some reflections on methodology and theory.

In the first paper, Reuther and Walter deal with a basic but critical aspect of metadata in the automated management of bibliographic information: personal name matching. Their methodology for testing different algorithms is empirical, thus similar to many other machine learning applications. Such empirical methodology is also present in the second paper by Sriharee and Senivongse, but in this case, the problem addressed is that of discovering Web Services with some given properties. This combines advanced architectural issues with metadata specifications.

A different kind of research is then used in Broeder and Wittenburg's paper, which provides the description of a set of metadata descriptors for 'Language Resources', and a set of associated tools and an infrastructure to use these. This is a second kind of typical research in metadata – providing a schema that extends or improves existing practice. However, the methodology to carry out such types of work is still not standardised, and follows mostly an ad hoc, goal-driven approach.

Grischenko's paper then provides an experimental evaluation for a metric of distance devised to cope with some characteristics of P2P networks. Here again the empirical dimension is present, but the artefact evaluated this time has to do with measurement, which is a major goal in metadata and semantics.

Then, Manouselis and Costopoulou present another schema description, this time for e-markets, with an expositive method that emphasises the lack of support of existing schemas for the domain addressed.

In the last paper, Khriyenko and Terziyan deal with extending RDF for representing context. This represents a totally different research standpoint, which is aimed at representation, but not specific or tied to a domain, but addressing a cross-cutting concern for metadata schemas.

It is clear from the reading of the papers in this issue that metadata and semantic research presents a wide diversity in several aspects, including formality, assessment techniques and domain-specificity. Further, studies range from the more technology-oriented to the more user-centred. This results in a 'mosaic' of non-homogeneous research papers that are vaguely identified as 'metadata' or 'semantics' related. One might wonder where are the demarcation points for such area of research, but in any case, it is the presence of metadata in some form that determines what can be included or not. This does not necessarily entail that every IJMSO paper needs to deal with metadata measurement or creation, since papers that deal with metadata schemas (and ontologies can be considered as metadata schemas as well) are of course also central to IJMSO readership.

Methodological issues in metadata research

The papers in this issue illustrate the different methodological standpoints possible in metadata research. Going a step further, we can state a number of methodological aspects that might help in characterising the space of methodologies applicable. Table 1 describes some of them in several 'axis'.

Table 1   Some methodological aspects of metadata research

Dimension	Examples
Formal/empirical	Formal studies on metadata might deal with computational frameworks for metadata transformation, the mathematical aspects of metadata languages or related aspects. Pure formal studies are not common in metadata, but they deal with some practical application aspects, and in many cases provide case studies, demonstrations or empirical data for the evaluation of metadata applications or practices
General-purpose/ domain-specific	Metadata research may address general purpose issues as how to better store or express general purpose metadata. But many of the papers in the field are tied to a particular domain/application, for which some highly specific schemas are applied or created from scratch
Technology-oriented/ human-social oriented	Some metadata research emphasises processing by software agents, while other focus on the better ways for humans to create useful metadata, or study actual human or social metadata creation and use practices
Managerial/technical	Metadata research may also deal with the managerial aspects of creating, searching and maintaining metadata, including considerations of trustworthiness in metadata and Knowledge Management issues

The diverse aspects collected in Table 1 may help researchers to frame their research into the research strands that are covered by IJMSO. It also provides some hints on directions of possible research regarding methodological aspects. For example, analysing and stating principles of metadata has been addressed elsewhere (Duval et al., 2002), but a systematic analysis on how these principles actually influence practice and existing metadata schemas is still missing. Such an analysis could help in sketching something like 'metadata engineering' as a discipline that covers metadata schema definition and assessment together with aspects of measurement and techno-social acceptability. These and others are ideas that I urge metadata researchers to deal with. They can be sure that IJMSO will warmly welcome discussion of such innovative ideas.

Acknowledgement

Reference

Duval, E., Hodgins, W., Sutton, S. and Weibel, S.L. (2002) 'Metadata principles and practicalities', *D-Lib Magazine*, Vol. 8, No 4, April.