Ontology Building as Practical Work: Lessons from CSCW

Yuwei Lin[1], Rob Procter[1], **Dave Randall**[2], John Rooksby[3] and Wes Sharrock[4]

[1]National Centre for e-Social Science, University of Manchester
[2]Department of Sociology, Manchester Metropolitan University
[3]Department of Computing, Lancaster University
[4]Department of Sociology, University of Manchester

Abstract

Ontologies are a key technology for the realisation of the e-Science aims of increasing the sharing and re-use of scientific data, and of greater collaboration in research. Ontology building can be thought of sociologically. By this we mean, the work undertaken and the problems and difficulties entailed can be understood in terms of the practices of knowledge workers and the practical nature of 'sorting things out'. It does appear that many of the problems in the work of ontology building carry a resemblance to problems in software engineering, particularly the engineering of cooperative systems. In this paper, we discuss research in the field of Computer Supported Cooperative Work (CSCW) that has focused on classification and which, we believe, throws some light on ontology building. We then introduce some early data from our own ethnographic studies of ontology building.

1. Introduction

The development of ontology-based computer systems for various kinds of knowledge work continues apace. In the context of e-Science, this, not entirely surprisingly, has focused in the main on the capture and re-use of scientific knowledge and process. Inevitably, this means that ontologies are being deployed into areas where knowledge of entities and their relationships is allied to procedural work, and the work of professionals (medicine and knowledge management work being obvious examples). These developments have come, however, with a recognition that increased complexity provides for a different order of problem. Dealing with procedural matters raises the spectre of 'exceptions'; dealing with professional endeavour with that of an acceptable allocation of function, and dealing with heterogeneous environments with that of the 'knowledge gap' between one community and another. In the ontology-building community, this is often expressed in terms of the size and scope of ontologies themselves, and the concomitant issue of modularization. The point we will make is that they can equally well be thought of *sociologically* – as issues to do with the assumptions that knowledge workers make when they do their work; the practical nature of their enquiries, and the assumptions they carry when they make them. If so, then there is considerable mileage in considering what kind of work sociologists can do in support of ontology-building.

Research in the field of Computer Supported Cooperative Work (CSCW) has addressed problems that, for us, resemble very much the problems faced by those engaged in the building of ontologies. In particular, we note a resemblance between ontology building and classification work. Rather than re-invent the wheel, it seems prudent to address this prior work. In this paper we discuss recent work by Martin, O'Neill, Randall and Rouncefield (2007) and an influential book by Bowker and Star (1999). We then introduce some of our early data from our ongoing studies of ontology building.

2. Ethnography and Classification

As social scientists working over a period of time in the field CSCW, we are interested in how the kind of ethnographic work we are typically engaged in can be related to ontology building. This is predicated on prior experience of the social sciences in the system design process, and notably the contribution that ethnography has made to it (see Hughes et al., 1994). The original complaint upon which CSCW was founded had to do with the inadequacies of mechanistic, top-down models in the system design process when confronted with a new order of complexity in patterns of use – specifically, when user communities worked on networked machinery, and where patterns of input and output were distributed across work environments in ways which were not well-known. It is our contention that, as ontology building moves out into the areas we

have described above, so the problems that have to be dealt with look increasingly like those experienced in CSCW. One of the solutions proffered has been the use of ethnographic methods to uncover detail about the business of work.

Ethnographies of work for CSCW have been conducted in a wide range of domains, including air traffic control, emergency services work, transport coordination, knowledge management regimes and so on (see e.g. Shapiro et al., 1991; Goodwin and Goodwin, 1993; Heath and Luff, 1992; Randall et al., 1996; Watts et al., 1996; Peterssen et al., 2002; Ackermann et al., 2003). Ethnographies tended for a time to be local, and small-scale (historically, emphasising aspects of control room work, or specific interactional problems associated with computer mediated communication) but in recent years, has turned to more complex domains, and to the knowledge work associated with them (see, for instance, Harper et al., 2000). Here, and it is as well to be clear about this, the emphasis has been on an alternative conception of knowledge, one which is quite distinct from the information-theoretic one embodied in much ontology-based work. The dominant metaphor in such research has become that of 'expertise sharing' (see Ackermann et al., 2003) and research has emphasised the ordinary practical ways in which knowledge or expertise is (or is not) shared across knowledge communities. The kind of ethnography done in CSCW, the way it is done, and the ways in which it is used to throw light on systems design is by no means the only way to go about doing ethnographic research. Our work is akin to that done in e-Science by Christine Hine (2003), for example, but should not be thought of as trying to achieve the same results for the same purposes.

We do not set out to critique assumptions about knowledge here, nor to privilege one set of beliefs about the nature of knowledge work above another. Our purpose is to identify how this problem of knowledge work and its behavioural dimensions resonates with some of the issues we have mentioned above. On the one hand, ontologies are a formal method of encapsulating knowledge and are intended to provide for the flexible re-use of data, and for rapid inferencing concerning data held. On the other, it is clearly the case that some ordinary, practical considerations go into the use and exchange of knowledge, and that some of that knowledge is not easily captured. In turn, it is a reasonable bet that an understanding of the latter will have some impact on the procedures for capturing the former, and the design of tools to support it.

Of course, and in principle, ontologies can come in any size, from the one-size-fits-all approach which we would associate with a realist paradigm, to the modularized, limited scope approach we might associate with a more pragmatist paradigm. It is the latter we can be more readily associated with here, however, we have nothing to say in this paper about the philosophical underpinnings of disputes between realists and pragmatists in the ontology community.

3. New Lamps for Old: The Allocation-of-Function Problem

To illustrate the problems that may well turn out to be relevant for the deployment of ontologies in complex organizational environments, we explore some recent work on classification schema as they are used in call-centre work (see Martin et al., 2007). Although nothing that looks like a fully-fledged ontology is in place in this work, there are embedded classification schema. In this paper, the authors, borrowing from Bowker and Star (1999), suggest that three dimensions are particularly important when looking at the degree to which the deployed classification systems turn out to be useful or not. As Martin et al. put it:

> "[we are] … looking at the design relevance of classification through examining its actual everyday, operational nature as it features in call centre work. We do so because it became increasingly evident to us, when looking over the material we were examining, that the standardization processes we were looking at were almost always classification processes – operators in the customer-facing work we were interested in were expected to see their encounters with customers not only in terms of a standardized procedure but also as 'types' of encounter, generated by schema embedded in machinery." (Martin et al., ibid)

In turn, they use the 'wide-ranging archaeology' of classification which Bowker and Star (1999) discuss, and issues that they raise in association with it:

> "… at the level of encoding, tools need to be sensitive to the working conditions of those encoding the data." (p159)

And

> "Imposed standards will produce work-arounds. Because imposed standards cannot account for every local contingency, users will tailor standardized forms, information systems, schedules and so forth to fit their needs…. When designing tools for distributed, organizational decision and policymaking, a detailed catalogue and analysis of such responses could become part of the designers' tool kit; incorporated in the system, it could point out styles of work arounds at the level of coding." (p159).

Bowker and Star, then, are concerned with a 'gap' they perceive between the 'formal' and the 'informal', and the ways in which we might better understand the relationship between the two. They raise 'challenges' for classification schema in terms of comparability, visibility and control.

3.1. Comparability

Comparability refers to a 'regularity in semantics and objects' (1999: 231) and thus pertains almost by definition to ontologies. What is important in this context, however, is the degree to which this stability is, in practice, obtainable. Most ontology-design hitherto has been aimed at relatively homogeneous communities, where underlying concepts (if not terminology) stand a good chance of being commonly held. Ontologies that have to serve more heterogeneous situations and purposes may turn out to be serving one user group more successfully than another – a problem that has been well-attested in the field of medical informatics. Problems of this kind will be compounded as and when ontologies are deployed across organizational boundaries. These might be used for managerial as well as professional purposes, or might be deployed in contexts where parties to the work have no knowledge of the categories that underpin an ontology (as, for instance, with customer or client-facing work). The potential mismatch between the 'requisite variety' of terms in an ontology, and the actual use of terms is shown by Martin et al. in the context of a help desk. They argue:

> "Organizationally, the system was intended as a repository of problem types and solutions to improve efficiency; to aid continuity of service through customer record, and to measure work (by recording call times). This data could then be used to potentially re-arrange work by distributing personnel or assigning personnel to dealing with subsets of problems. However, some mismatches appeared between the imagined purposes of the system and the actual mechanics of classifying… [for instance] .. a large amount of calls were classified as 'other' or 'miscellaneous')",

And go on to show how:

> "… many of the 100 or so categories of problem were rarely used. The types of problem listed under a category could vary quite greatly. One 'password problem' could easily be different from another, while 'general problems' showed massive variance. Within the flow of work the operators viewed problem classification as an inexact activity that was carried out under time pressure rather than an accurate portrayal of the work. The written commentaries in associated fields in fact often provided a clearer memory of the work undertaken than the classification scheme itself, offering the possibility to facilitate continuity of service through an audit trail of customer-organization interaction. They provided a way into a problem and were commonly backed up with the recollections of staff."

Desk operators here utilized the system as a resource for standardizing practice, but actually managed the business of consistency through observing and talking with one another, supplementing 'the records' with occasioned collaboration. The records did not speak for themselves but were made sensible through cooperative techniques. Similar conclusions have been reached in other work which shows how distinct problems of classification occur in customer-facing work (see Harper et al., 2000) It is one thing for a professional knowledge worker to understand the uses to which an ontology can be put; it is quite another for that person to interact with 'outsiders' as they are doing so, and this relates to the notion of 'visibility'.

3.2. Visibility

The second of the challenges to classification recognised by Bowker and Star is visibility. 'Invisible' areas of work are "by definition unclassifiable except as the residual category: 'other'" (1999: 231) 'Invisible' work here refers to those informal practices which do not, in themselves, constitute part of the ontology, but which may nonetheless be critical to how an ontology is constructed. For instance, it would

be central to questions as, firstly, how explicit and complete should the categories in any classification scheme be, and hence to decisions about appropriate modularity for ontologies. For Bowker and Star there is an inevitable trade-off between comparability and visibility, insofar as comparability allows for use across a variety of settings, but risks an increasing degree of inappropriateness for each local setting where it is used. They make the point that the more comparable (aimed at use across a number of settings) embedded classification schemes are, the less visible the work that goes into maintaining them will be and the less they will fit 'local' arrangements. This, we think, is critical to decisions about the scope of ontologies, and, more importantly, raises the question of where our knowledge about these different settings – knowledge that would enable us to make appropriate decisions – is going to come from. Equally, and as pointed out by Martin et al., 'translating and mediating' work is frequently invisible, but necessary. In other words, in complex organizational environments, and dare we suggest even in professional knowledge work, we cannot presume that all users are equally adept at using ontologies. Put simply, what is being done with a classification system sometimes has to be explained to other people.

This, of course, implicates the interface – certainly at the user interface level and possibly also at the API (application programming interface) level as well. It is one thing for the knowledge worker to understand what is in front of him/her in OWL or in an editor such as Protégé-OWL or Swoop, but quite another for he or she to use it to explain to others what work needs to be done. Whilst it is becoming common to visualise ontologies in various ways, particularly as a graph, we would suggest this problem can be addressed far more extensively. The work of Buckingham-Shum and his collaborators on rationale and argumentation has already served to point e-Science in this direction. In their work, attention has been paid to integrated and interoperable technologies for presenting and manipulating data but also to "the art and craft … to know how to use the tools well enough that they are constructively disruptive, delivering immediate value to those using them, as well as supporting longer term memory" (Buckingham Shum et al., 2006, p129). If ontologies are to be deployed in heterogeneous domains where interfaces need to be shared amongst people with varying degrees of expertise then a great deal of thought needs to go into how classifications are represented.

3.3. Control

Regarding control, Bowker and Star argue that,

> "… any prescription contains some amount of control to be exercised by the user, be it as small as in the most Taylorist factory or prison or as large as the most privileged artists' retreat ….. the managerial trick is to measure the degree of control required to get the job done well, for most people, most of the time."

Again, and especially where ontologies are to be extended into procedural requirements, this would seem to be critically important, because users of any given ontology will themselves make decisions about its use in the light of their need to get the job done well. As Martin et al. suggest, based on their data on call-centre work-problems often occur when organizations attempt to switch the expertise from the operator to the system. They say:

> "Control is central to the design of classification–based systems, in our view, because it is central to the problem of the residual. Skilful decision-making work might turn out to be a means by which the residual can be avoided or alternatively residual categories might create problems for that work. In other words, work can be, and may have to be, done in order to translate what would remain residual into one category or another … Rehearsals of rules and their application, or as they are more commonly termed in CSCW research, plans and situated actions, are precisely examinations of this control problem. As already noted, however, Bowker and Star (op cit) … point to the practicality of issues surrounding degree of control. As they say, "This balance can never be fully resolved (as novices and strangers are always entering the field of work) …"" (op cit: 232)

If Bowker and Star are right in that these issues are central to any classification system, and if we are right that they pertain as much to ontology-building as to any other kind of classification scheme, then it would seem that there are good reasons for examining the work that goes into the construction of ontologies – classification practices, if you will – because that work will ramify in the development of ontologies and of systems to support their

development. Their success or failure when deployed will not only be a matter of their internal consistency but also the degree to which they meet organizational requirements. There may be a number of dimensions to this, and below we sketch out what some of them may be, based on our own observations of ontology-building work conducted over a six month period.

4. An Ethnography of Ontology Building

We now turn to a different set of issues, to do with the process of building of ontologies itself. We mentioned above the idea that all human practices can be viewed as ordinary, practical matters. We have already seen aspects of this in our own ongoing ethnographic studies of ontology-building work. Drawing on findings from these studies, what we do below is discuss aspects of ontology building work-in-practice that seem to have ramifications for supporting community based ontology building. This discussion covers issues of identifying purpose, rationale recording and timeliness.

Our first data is taken from an ontology building course on a Masters' degree. The following is an example of the kind of instruction students are given concerning how they might proceed with the business of building an ontology:

> *"The first thing you have to do is establish the purpose. Without a clear purpose, there can be no scope, no requirements and no evaluation. It's hard to constrain the problem without something in mind, even though you might want to re-use it."*

A tension between re-use and specific purpose is evident, and we will refer to it again. In reference to term-collection, a necessary initial step in ontology building work, we hear the following:

> *"Organize them informally, paraphrase and clarify them to produce informal concept definitions … paraphrasing is really important … how will we know what you're trying to do if you don't make notes … photographs are good too … I mean it, if you have a digital camera or a mobile phone, take photos of the way you organise the cards … felt boards are useful tools for organizing things …"*

What is interesting about this is the way in which the business of ontology building becomes the business of understanding and noting one's own rationale for making the decisions one makes, and the use of some very prosaic techniques – photographs, card-sorting, felt boards, and so on – to do so. This is linked to an explanation of how one comes to decide that one's hierarchy of concepts might be done this way instead of that:

> *"Card sorting is as good a way as any … the metaphor is highlighting … here's a list and you want to group things together … you have to make decisions about how you're going to do that … this is sometimes called, 'laddering' … group things and then ask why … what do they have in common … what are their parents and siblings? What other siblings might there be … for example, if you're doing an ontology of children's animals, there are bacteria and fungi, but do you want to include them now? You're going to get lost in the trees, and it's very easy to lose sight of the woods …"*

Reference is made to disagreement, because one of the evident features of ontology building is lengthy disagreement, and the need to resolve it at some point.

> *"Without a paraphrase you can't disagree on why we did something … what can we say about all members of a class? … all of this does some of that or all of this … these are the only two constraints we've got."*

Equally, the practical issue of recording decision-making, remembering rationale, and so on, will become a great deal more important as ontologies are designed for more heterogeneous communities, in part because the actual process of building will become a more distributed affair. If we look at the kinds of technique deployed above for the recording of rationales, disagreements, etc., then it is quite evident that they will have to be replicated in some other way if distributed ontology building is to be possible. These issues tie in closely with the kinds of debate that inform the ontology-building community today – debates about problems of scale, scope, detail and usefulness. Whether or not, as the realists would have it, an ontology of everything is possible, as a matter of practical contingency important decisions have to be made about the number of classes to be identified, their properties and relations. That

is, the ontology needs to classify persons and events in such a way that it covers all relevant possibilities (including the rare exceptions), and discriminates them sufficiently such that the consequences of different events and behaviours can be identified and dealt with. How we begin to identify what "relevant" might mean here is entirely non-trivial.

Closely related to this is an issue which we can describe for convenience as temporal. That is, understanding the processes of development across time may be revealing. The kinds of problem this may entail can be seen in the following:

> *"Of course, a feature of microarray technology is that massive amounts of data are produced …petabytes … there are innumerable difficulties attached to this, not least that across the range of proteomics; transcriptonomoics; metabolonomics, etc., there will be a range of different experimental methods with different approaches to normalization; different ranges tested, etc. … even so, the experimental metadata I get is all the same …. the problem is that biologists don't see sufficient gain in inputting this data themselves- it often remains invisible. An ontology which provides some of this data would minimize 'cost' and maximize 'benefit'. The trouble is, it involves an awful lot of drudge work, especially in respect of coming up with definitions that everyone agrees with. It would have a particular benefit, in that it might provide for 'environmental' information that some biologists would not otherwise think to provide because they don't need it."*

It is not an especially new observation that it is difficult to get people interested in development when they do not see any immediate practical benefit for themselves. One aspect of the heterogeneous community we have been speaking about, then, is a disparate set of interests. The point is that these varying interests have all sorts of consequences for the way in which any proposed ontology might be structured. Hence:

> *"Some of what we want to do is just controlled vocabulary, but some of it is to do with other issues, like granularity … problems we have to deal with include the fact that habitats are not discreet, they blend into each other. They can be described in many ways."*

Related to this is the way in which these different interests coalesce around the practical business of 'getting involved'. This turns out to be hugely problematic. Much of our data concerns the way in which it is difficult to get involvement at the point where the builder needs it, but easy when others in the community see it in terms of their purposes. We would suggest, however, that where ontologies are to be deployed in complex environments, a great deal more exploration of what those purposes might be is needed. The reasons for this quickly become obvious:

> *"We're going to have to consult a lot over terms, but maybe will have to legislate to some degree. We're going to have to allow people to tag stuff up. To some extent, at least to begin with, there are some agreed boundaries … tropical rainforests have agreed attributes … and these would be largely unchallenged. There are, however, about twenty different types of grassland within that habitat, defined not only by the type of grass but also maybe by the fauna as well. And then there's all this to do with temperature, rainfall, etc, etc. We wouldn't want to include geospatial information. Habitats grow and shrink but we don't need that …"*

The issue of 'consultation over terms' again has practical consequences for ontology building, for there is controversy over when and how that should take place. On the one hand, as one respondent observed, 'GO [Genome Ontology] started from a use case … that's why it was so successful …' In other words, early engagement with users can be seen as desirable. On the other:

> *"It'll just be a controlled vocabulary to begin with … with most successful ontologies, the complexity came later. Sometimes you feel like a lawyer, finding descriptions that no-ones going to disagree with …"*

It seems to be a common experience among ontology builders that getting any kind of community agreement over terms in the earliest stages is extraordinarily difficult. It can result in the kind of one-size-fits-all ontology we refer to. Thus:

M: OBI is such a big beast because people felt it had to be all-encompassing and centrally managed …'
B: 'A big semantic cricket bat'
M: 'There was no modularity …'

In contrast, it seems that success sometimes comes from a very slow and careful form of user engagement:

"We're trying to keep it quiet … if we try to develop by committee we're not going to get anywhere for a long time … but sooner or later we need to put something out there into the community …"

5. Discussion

It would seem that problems faced in ontology building carry a strong resemblance to those identified in Computer Supported Cooperative Work (CSCW). Ethnographic research has been of benefit in CSCW in identifying and understanding these problems. However, we are not suggesting that ethnography in itself offers solutions, and we do not offer definitive conclusions in this paper about the most effective ways to develop ontologies in the community which they are meant to serve. Our work here is focused simply upon elucidating the problem of ontology building.

The issues of how to go about building, maintaining and reusing ontologies have been a longstanding concern, and it seems researchers in this area have long understood the need to focus on 'how to do it' hand in hand with 'how it (really) gets done' (e.g., see Gruber, 1993). Technologies and techniques to support the work of ontology builders are being developed with increasing pace and sophistication (not least the ontology development environments such as Protégé-OWL). Our belief is that this is an area in which our ethnographic work can be of benefit.

In this paper, we argue the first step to designing solutions is to identify problems. The problem-set we have recognised has to do with a range of issues such as rationale-recording; identifying purpose; the timeliness of community involvement, and so on. All seem to us to be deeply relevant to the success of ontology building processes, and depend on accurate and adequate knowledge of how those processes are currently managed.

References

Ackermann, M., Pipek, V. and Wulf, V. (2003). Sharing Expertise: Beyond Knowledge Managment, Cambridge, Mass., MIT Press

Bowker, J. and Star, S. (1999) Sorting things Out: Classification and Its Consequences. Cambridge MA: MIT Press.

Buckingham Shum, S.J., Selvin, A.M., Sierhuis, M., Conklin, J., Haley, C.B., Nuseibeh, B. (2006). Hypermedia Support for Argumentation-Based Rationale: 15 Years on from gIBIS and QOC. In Dutoit, A.H., McCall, R., Mistrik, I., and Paech, B. (eds.) *Rationale Management in Software Engineering.* Berlin: Springer.

Goodwin, C. and M. Goodwin (1993). Formulating planes: Seeing as a situated activity. Communication and cognition at work. Y. Engeström and D. Middleton. Cambridge University Press.

Groth, K. and J. Bowers (2001). On finding things out: situating organisational knowledge in CSCW. In Proceedings of the European Conference on Computer Supported Cooperative Work, Bonn, September.

Gruber, T (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies* Vol. 43, Issues 5-6, November 1995, p. 907-928.

Harper, R., Randall, D. and Rouncefield, M., 2000. Organisational Change and Retail Finance: An Ethnographic Perspective, London, Routledge.

Heath, C. and P. Luff (1992). Collaboration and control: Crisis management and multimedia technology in London Underground control rooms. *Journal of Computer Supported Cooperative Work* 1(1-2): 69-94.

Hine, C. (2003). Systematics as cyberscience: the role of ICTs in the working practices of taxonomy. Paper presented at the Oxford Internet Institute "Information, Communication and Society" symposium, 17-20 September 2003, University of Oxford, UK.

Hughes, J., King, V., Rodden, T., Anderson, H. (1994). Moving out from the control room: ethnography in system design. In Proceedings of the 1994 ACM conference on Computer supported cooperative work, p. 429-439.

Martin, D., O'Neill, J., Randall, D. and Rouncefield, M. (2007). How Can I Help You?

Call Centres, Classification Work and Coordination. Accepted for Publication in the *Journal of Computer Supported Cooperative Work*.

Pettersson, M., Randall, D. and Helgeson, B. (2002). Ambiguities, awareness and economy: a study of emergency service work. In Proceedings of the 2002 ACM conference on Computer supported cooperative work, New Orleans, Louisiana, USA.

Randall, D., O'Brien, J., Rouncefield, M. and Hughes, J.A. (1996). Organzational Memory and CSCW: Supporting the 'Mavis' Phenomenon. In Proceedings of the Sixth Australian Conference on HCI (OzCHI '96).

Shapiro, D.Z., Hughes, J.A., Randall, D. and Harper, R.R. (1991). Visual Re-representation of Database Information: The Flight Strip in Air Traffic Control. In Proceedings of the 10th Interdisciplinary Workshop on Informatics and Psychology: Cognitive Aspects of Visual Language and Visual Interfaces, Scharding, Austria, May.

Watts, J. C., D. D. Woods, J. M. Corban and E. S. Pattersson (1996). Voice Loops as Cooperative Aids in Space Shuttle Mission Control. In Proceedings of the Conference on Computer Supported Cooperative Work. Cambridge MA USA, ACM press: 48-56.