

Modeling Social Information Skills

Daniel Memmi

Dépt d'Informatique
UQAM
Montreal, QC, Canada
memmi.daniel@uqam.ca

Abstract: in a modern economy, the most important resource consists in human talent: competent, knowledgeable people. Locating the right person for the task is often a prerequisite to complex problem-solving, and experienced professionals possess the social skills required to find appropriate human expertise. These skills can be reproduced more and more with specific computer software, an approach defining the new field of social information retrieval. We will analyze the social skills involved and show how to model them on computer. Current methods will be described, notably information retrieval techniques and social network theory. A generic architecture and its functions will be outlined and compared with recent work. We will try in this way to estimate the perspectives of this recent domain.

Keywords: social information retrieval, information retrieval, social networks, social skills, collaborative systems, knowledge representation.

1 Introduction

In a knowledge-based economy, the main resource is no longer raw materials, physical infrastructure or even formalized knowledge. Today's most important resource is skilled people, their expertise, problem-solving abilities and creativity (Drucker, 1992). What matters most nowadays is to locate and make use of the right persons for a particular task, and experienced professionals usually possess the appropriate social skills for this purpose. They know how to find experts and how to evaluate their competence and reliability.

Because of the rapid spread of electronic communication technology, social activity leaves many traces online and computers can help us find the right people. In the same way that we would use various social clues to find our way in society, we may (and often do) use electronic information to locate competent people. This quest can be automated to some extent with the help of computer techniques, and social information retrieval is now an active research area (Goh & Foo, 2007). Analyzing how human beings retrieve social information is a good start toward modeling social skills, but we must also develop adequate software (Memmi, 2007).

Among potentially relevant techniques, classical information retrieval (dealing with textual documents) stands out, as this mature field has elaborated extensive

representation, indexing and search techniques. This domain is very useful because document retrieval methods are often applicable to person retrieval as well. Social network theory is another important field, which shows how to represent and exploit social information formalized as graphs, a well-understood and efficient data structure. Social networks are therefore attracting more and more interest.

In this text, we will first study human social retrieval skills, and we will then see to what extent they can be modeled with currently available techniques. Information retrieval and social network analysis will constitute the most obvious methods, but other approaches should also be considered. An appropriate generic software architecture will be proposed and compared with recent research work. We will try at the same time to evaluate the perspectives of this new field.

2 Basic social skills

Some basic skills are required for social information handling. We will analyze the social skills involved before trying to come up with computer models. When embarking upon some new problem-solving assignment, the first task is actually two-fold: retrieve relevant documents and locate appropriate people. Textual documents can be found with information retrieval techniques, but socially-situated information, which is associated with specific persons, requires considering the social context and usually involves some dialogue. Various forms of collaboration are thus necessary for most open-ended tasks.

It will be useful at this point to standardize our vocabulary. We propose using the term *request* (rather than the more specific document *query*) for some general inquiry about either documents or persons, *requester* for the person making a request, *document* for any item containing explicit information, *expert* for a presumably competent person, *retrieval* for the process of locating some document or person, *referral* for the reference by one person to the next in a chain of referrals from requester to expert. We will use masculine pronouns (*he, his*) in a generic manner regardless of actual gender.

What are then the fundamental social skills needed for collaborative work? Without too much regard for technical implementation for the time being, we can rank skills from the easiest to the most difficult to model. They are fairly general and could apply equally well to the 17th century Canadian fur-trader trying to secure reliable Indian partners or to the 21st century lawyer looking for an expert in some fine point of commercial law. These skills are often more or less intermingled in practice, but we will distinguish them here for the sake of clarity:

1) expert retrieval

The first task is to find a person competent enough to answer questions, provide information, suggest documents, solve problems, come up with fresh ideas... Locating somebody with the pertinent expertise is often a prerequisite to performing the high-level, knowledge-rich, open-ended tasks typical of a modern economy. Even

experienced professionals rarely have all the expertise required for today's increasing variety of tasks, whether in law, science, engineering, medicine or business.

Expert retrieval can usually be accomplished by following two or three steps in a referral chain: ask an acquaintance (close colleague for instance), who seems somehow familiar with the subject of the request, for the name of a more qualified person, and so on... This strategy has proven to be fairly efficient and can be modeled on computer. Alternatively, one may consult an index or database of some sort, a process which is also easily automated (provided such a database does exist, which is not always the case).

2) checking competence

Even when an expert has been found, one cannot be sure of his competence, all the more so as the requester might not know enough about the subject to evaluate an expert's level of ability. Before acting on expert advice, it might be a good idea to try to check the validity of such advice! Documents (if any) associated with the expert provide a first indication, but there are also social indicators of competence.

Fortunately, the very process of expert retrieval furnishes clues as to his reliability. The referral chaining process indicates the social integration of each link in the chain: people refer to persons they trust. If the chain is not too long, there is a good chance that the final expert belongs to the same world as the requester and complies with similar norms of competence. Moreover, there are various social indicators of competence which can be consulted: diplomas, professional associations, prizes and honors, overall reputation...

3) gauging responsiveness

Knowledgeable experts are often busy people, and the name of an expert is not of much use if one cannot be reasonably confident that he will reply to a request. In social life, various schemes are used to estimate the good will of experts, and requests often begin with a reminder of common acquaintances. Roughly speaking, the closer and the stronger the relationship between requester and expert, the better the chance that questions will be answered willingly.

This is usually the case within official organizations (business firm, university, ministry...) provided the organization is not too big; staying within a department or service might then be a safer bet. But networks of social contacts, frequent relations and similar interests are also good indicators of responsiveness, even when they cut across organizational affiliations. Asking for an explicit recommendation from a common acquaintance can also be useful.

4) evaluating trustworthiness

Another problem is that an expert might not be trustworthy or benevolent. He could lie, fail to mention relevant information, or just not do his best to answer a request for help. Reasons could range from lack of time or interest to feelings that the requester is potentially in competition with the expert. Also "knowledge is power" and many people are reluctant to share hard-acquired expertise without due return.

It is in fact a crucial social skill to be able to evaluate and to motivate the good will of others. This usually involves some kind of barter relation, whether explicit or implicit. One can trust someone when a common framework ensures a regular exchange of services, so that it is in the interest of each partner in a relationship to respect the interests of others, in the future as well as in the past. Identifying such patterns of exchange will help estimate the reliability of experts.

5) dialogue abilities

When all the preceding conditions have been met, it is still not certain that interlocutors will be able to communicate effectively. Understanding the other person's point of view, technical jargon and particular needs is crucial for a successful dialogue, whether on the part of the expert or the requester. This is best taught by experience, and some experts prove better than others at dealing with vague, inaccurate or seemingly pointless questions.

But it would also be helpful to help requesters formulate more precise and relevant questions. Establishing a common vocabulary and a shared conceptual framework is actually the first task to perform when trying to achieve collaborative work. This requires intellectual flexibility and can prove very frustrating, but technical help is possible, from a traditional thesaurus to recent ontology editing software.

6) negotiating skills

The more difficult the task to perform, the more effort will be required of the expert to satisfy complicated requests. This must be understood by the requester, who should be ready to offer something in exchange for the expert's time and effort. Identifying (and formulating) common interests, gratifying an expert's personal goals, playing on a pattern of recurrent exchanges will be useful in order to motivate the expert. The implicit threat of retaliation in case of non-compliance is usually best left unsaid, but also belongs to the arsenal of social skills.

In other words, one should always consider that "what's in it for me" or "what if I don't comply" are likely to be on the other person's mind. Successful negotiations try to answer these questions. Business negotiators, politicians, diplomats, union officials are usually good at this sort of give and take. But the problem now is that such considerations are often implicit and hard to formalize with presently available techniques. Multi-agent modeling is promising, but still difficult to use in practice.

7) problem-solving

Once the appropriate information or expertise has been located, a real task still requires problem-solving skills. They include the ability to formulate (or reformulate) a problem in such a way that it can be tackled with the resources at hand, as well as the application of pertinent techniques and the capacity to carry out an implementation. Although research in Artificial Intelligence has addressed these issues to some extent, this is probably going beyond what we can expect to model with present software techniques.

Establishing an adequate conceptual framework is often a prerequisite for problem-solving. This largely explains the present interest in ontologies, and work in this area may be a source of technical ideas. But this is only a partial requirement, and we still seem far from being able to mechanize problem-solving completely.

3 Currently available techniques

Two domains stand out as a source of potentially useful techniques to model social skills: information retrieval and social networks. These are both mature fields in which decades of work have come up with elaborate conceptualizations, effective data structures and practical algorithms. They both offer a toolbox of methods to tackle the social skills we have outlined in the preceding section. We will consider these two domains in turn.

Another domain relevant to social modeling could be called collaborative concept formation, i.e. the collaborative elaboration of a shared conceptualization. This is still a recent and rather diffuse research area which has not reached intellectual or social maturity, and which does not even seem to have a generally recognized name as yet. But interesting work is being done in this area, and we will try to describe some of the directions currently pursued.

So here are the three main domains pertinent to modeling social retrieval skills:

1) information retrieval

The primary concern of information retrieval has consisted in indexing and locating documents as efficiently as possible (Salton & McGill, 1983; Baeza-Yates & Ribeiro-Neto, 1999; Manning & Schütze, 1999). This field shows how to represent items of interest and how to compare them so as to estimate their semantic proximity. The typical application consists in retrieving relevant documents, given a short query containing a few words. For instance, everybody is now familiar with Web search engines which make it possible to quickly retrieve relevant Web sites among millions of pages.

Though quite a few technical variants have been proposed, the basic idea is to represent a document by a vector of significant words (usually the most frequent and distinctive terms). Document vectors can then be compared in vector space (using variations on the dot product) to decide which are closer to a query or any topic of interest. We will leave aside for now how this lexical representation is extracted from documents, the indexing techniques designed for efficient retrieval, and the precise comparison algorithms to evaluate semantic proximity. The main point is that a difficult semantic problem (evaluating similarity of content) has been transformed into simple and rapid numerical computation.

The items of interest are typically textual documents (e.g. text files or Web pages) but they could be anything as long as some representation can be devised for retrieval purposes. Images, sound files, video clips can thus be indexed and retrieved (although this is usually done by working out some textual representation of non-

textual items). Similarly, one can represent persons in such a way that retrieval becomes easily feasible.

Because so many people now use computers for text processing, electronic messaging or Web browsing, it has become possible to compute a characteristic individual profile for regular computer users. This can be done in various ways, notably by extracting index terms from files produced or archived, Web bookmarks or browsing patterns, e-mail messages sent and answered. One should note, however, that an individual may have several topics of interest, which should not be merged together into a unique representation. But once a representative profile is available for each user, it becomes as easy to find the corresponding person as to retrieve documents from a simple query.

In short, whatever the technical details, if one can devise some distinctive representation of individual people, person retrieval can be achieved with classical information retrieval methods.

2) social networks

Social network theory is a branch of sociology, considering social groups as a system of relations (Degenne & Forsé, 1994; Wasserman & Faust, 1994; Lazega, 1998). This structural approach, deliberately trading off the complexity of social life for a simplified, more abstract representation, has become totally formalized, and hence amenable to computer modeling. Social actors and social relations are represented as a graph, nodes standing for actors (individuals or groups considered as units) and links for relations between actors. The adjacency matrix corresponding to this graph can then be used to compute all kinds of operations on the network, and notably the extraction of significant subgroups.

This approach shows how to represent and exploit the position of a social actor within a network of professional or personal relationships (Burt, 1992). Sociologists working in this domain have mostly been concerned with social prestige, power relations or group dynamics. For example highly-connected actors usually control more social resources than marginal nodes. But network structure may also be used to understand the diffusion of innovations and the spread of epidemics. In the context of social information retrieval, social networks can be exploited to locate an expert, check social integration, evaluate trustworthiness, etc.

To build a structural graph, one must first collect the data necessary to establish relations between individual actors. A link will be drawn between two nodes if a relationship is attested between the two corresponding actors. Structural sociologists must traditionally resort to lengthy investigations to collect the appropriate data, but here again widespread computer usage has drastically changed the problem. Structural data can now be extracted or inferred from computer traces of social activities, such as e-mail messaging or co-authoring articles.

Once the graph has been built, many operations can be performed on the corresponding matrix. Some are global measures like graph density (the proportion of actual links) or clustering (a measure of graph structuration). There are also a variety of methods for extracting highly-connected subgroups from the whole graph. A clique is a typical subgroup: all the nodes of a clique are interconnected. But most

real social groups are not totally connected, and other types of groups of been defined with looser standards of connectedness (e.g. n-clique or k-plex). For instance an n-clique is a set of nodes which might not be directly connected, but which are mutually accessible by a path of n steps (usually 2).

In fact the exact subgroups extracted will depend on the algorithm chosen, and some expertise is necessary to interpret the results, but a clearly structured graph should yield interpretable subgroups. These substructures show which groups individual actors belong to, and their social position within the group. This is obviously relevant to evaluate the social integration of a person, and his reputation or trustworthiness when looking for a reliable expert. And of course the graph itself can be used to locate an expert by following social links.

In short, social network theory (also called network analysis) provides formal representations and effective computer algorithms to investigate and exploit the structure of social networks.

3) collaborative concept formation

This is not yet a coherent domain, but rather a collection of various research attempts (e.g. Voss et al., 2000; Nakata, 2001; Zacklad, 2006). What they have in common is the idea that a shared conceptual framework could be developed through collaborative efforts. Several mechanisms have been proposed, but ensuring the coherence of the resulting conceptual structure remains a problem. This domain can be linked to recent work in ontology development (an ontology being a public formalized conceptualization) but a collaborative approach makes it so much harder to ensure conceptual coherence.

Research in this area is still tentative. Bottom-up (unsupervised) classification of common documents or thematic tags is one possible approach to produce a shared conceptual hierarchy. Document clustering seems to be the preferred approach, but resulting conceptual structures are often difficult to interpret and may have to be edited for clarity. Tagging documents (or any items of interest) by individual members of a group may help other members retrieve relevant items and may lead to a common terminology, but does not guarantee the formation of a coherent conceptual map.

On the other hand, research issues in this area are becoming clearer and practical technical advances are most probably to be expected in the near future.

Lastly, we should mention recent work in modeling of sociological phenomena with multi-agent systems (e.g. Wooldridge, 2002; Sun, 2001; 2006). This branch of Artificial Intelligence is largely experimental, although it also borrows from more formal work in game theory or dynamical systems. Collective decision-making, negotiation processes, or coalition formation can be modeled with an increasing level of detail, but most research in this area deals with small-scale, simplified problems and is not ready yet for practical applications.

To sum up, real collaborative problem-solving still appears too complex for modeling. So higher-level abilities such as negotiating or dialogue skills remain basically beyond the scope of automation. If one stops at expert retrieval, however,

negotiating or problem-solving skills are probably not necessary: they really belong to the next stage of collaborative work, and could be left to human beings or to future research.

4 System architecture

After this survey of problems and techniques, let us now consider what might be an appropriate software architecture for social information retrieval systems. Given the number of issues involved and the fact that research in this area has not settled down yet, a generic architecture would allow experimenting with various representations and algorithms. As information retrieval and social networks appear to be the two most relevant technical domains, such a generic system should accommodate these two areas at least.

This architecture could then serve both as a software platform for developing practical systems, and as a modeling framework able to reformulate various research attempts found in the literature. Studying what diverse systems have in common gives us a good picture of the basic requirements for social information retrieval.

We need a system containing enough data for both information retrieval and social network techniques to operate. The obvious choice is to build the system around a data structure exhibiting social links as well as textual information. Such a structure is rather similar to the structure of the Web: a network of links with information at the nodes. In fact, this is but one more example of the wider problem of modeling various types of networks (Lord, 2007).

A social information retrieval system would thus operate on a global data structure, consisting of a network of individual nodes, with (mostly textual) information associated with each node. The data structure must allow the system to retrieve any node in two different ways:

- by following connection links within the network
- by using information associated with the node

The system will then be able to operate in two basic modes: structural retrieval (following network links) or information retrieval (using individual node information). The two main elements of the central data structure being the links and the nodes, these should be described in more detail:

1) network structure

a social network is basically a graph of links, which can be represented by a square matrix (a triangular matrix is sufficient if links are symmetrical). This matrix is usually sparse, however, and a straightforward matrix implementation by arrays is wasteful in terms of storage. More efficient representations (e.g. listing links explicitly) should be considered for larger networks (at the cost of more processing time to explore the graph). But this is basic computer science, and social network software packages (e.g. Ucinet) offer adequate graph-handling tools. The nature of

links may be left unspecified, but each link may also be labeled with the type of relationship (collaboration, acquaintance...).

2) node structure

a certain amount of information is needed at each node to represent the profile associated with each individual. An array of pointers (for the whole node set) can give access to a block of information per node. This block can be implemented as a frame with specific slots, and may be seen as a kind of extended business card. There could be slots notably for the social function, various interests, professional associations, and the subgroups a person belongs to (if they have already been computed). This information may be symbolic (name, label) or numerical (number, pointer, vector), depending on the particular slot, but it should be easy to use and maintain.

This enriched network makes it easy to operate in either structural retrieval or information retrieval modes. For information retrieval purposes, node information should be indexed whenever possible, to avoid having to search the graph. But of course the very logic of social retrieval means it will also be necessary and revealing at times to follow graph links in order to find a particular node.

In such a system, data structures and algorithms are kept separate, which is the usual software engineering practice. But one could go further and propose an "active network", with computation taking place at the nodes. Many variants are possible, among which could be distinguished two main options according to the type of computation associated with nodes: each node has exactly the same computational ability, or different nodes may have different computational capacities.

For example some nodes might be considered more intelligent or more active than others, or they might have better resources at their disposal. Actual computation could be performed by a general controller running on a central server, or it could be distributed among many servers. If a different computer is devoted to each node, we have a fully-fledged peer-to-peer system with distributed data and potentially huge computing power. The computing abilities of each node might be expressed as declarative rules for easier maintenance.

Such an active network would be very similar to a multi-agent system, which might be useful for sociological modeling. Each node with its own computational abilities could represent an active social actor, and the whole network could model complex negotiations, coalition formation, problem-solving, and so on. This latter approach, however, appears to be farther away from our initial concern: retrieval with simple, practical techniques. It would be best left for higher-level modeling efforts, in order to investigate more complex social behavior.

We do not as yet envision a specific architecture for collaborative concept formation, as this research area is still too diverse. Still, the social network structure may help identify subgroups within which shared concepts might be useful, and semantic tags could be associated with individual nodes as well as with documents. But for practical applications a generic architecture must be kept as simple and open as possible, in order to accommodate various representations and techniques.

5 Typical work

We will not attempt an exhaustive survey of the field, but will now describe a few relevant examples highlighting possible dimensions of this research domain. So here are some typical systems, which could be easily reproduced with our generic architecture:

1) expert retrieval

Referral Web was a seminal project for expert retrieval using social links (Kautz et al., 1997). In this system, social links are extracted automatically from the Web by observing the co-occurrence of personal names in Web pages. For example co-authors of publications in bibliographies, members of university departments or common participants in newsgroups would be considered linked by a social relationship.

Starting from a given name or a small community, the Web crawler takes a day or two to generate a network with hundreds of nodes. This network can then be used to locate experts on a particular topic by following social links. The initial query might have to be compared with personal information on individual homepages to check if an expert is relevant to the topic, but the system emphasizes referral through the social graph.

A major advantage of this approach is that users do not have to provide any information and social network members may not even be registered with the system. The social network is built from publicly available information only and raises no privacy concerns. Referral Web does not seem to try to evaluate the level of expertise or the social integration of the experts it finds, but this was an early research project.

In short, Referral Web is a typical system for expert retrieval through social links.

2) mixed retrieval

Human Links is a good example of a mixed system combining expert retrieval with document retrieval (Memmi & Nérot, 2003). This is distributed search engine exploiting the collective expertise of users thanks to a peer-to-peer architecture. Once registered, a user becomes a node in the peer-to-peer network and will tacitly collaborate in the search for information.

The system automatically computes a specific profile for each user (from the user's Web bookmarks, but other information might be used as well). The similarity between profiles defines a virtual network and virtual communities, but no permanent social network is actually constructed. Users are considered close if their profiles are similar, but this could also be modeled by building an explicit social graph.

When a user submits a query, a request is propagated through the peer-to-peer network to locate potentially relevant experts by comparing the query with user profiles. When an expert is found, the system may then function in two different modes: ask the expert whether he is willing to help, or use his profile to expand the initial query and resume the search. The software can thus be used to retrieve documents as well as human experts.

A drawback of this approach is that users must first register, causing a cold-start problem: reaching a critical mass is necessary for the system to become really useful. But Human Links was fully developed as a practical collaborative search engine.

3) social document retrieval

Recent work shows how to use social networks for information retrieval purposes (Kirsch et al., 2006). The first step is to extract a social graph from a mailing-list archive: e-mail users are linked with their correspondents. The resulting graph is a typical social network with a small diameter (a small world) and a high level of clustering. A social network might also be extracted from other types of information (such as co-authoring) and the basic idea would still be applicable.

Starting from a query, the aim is to retrieve relevant documents using social as well as textual information. Documents are first retrieved with classical information retrieval techniques, comparing the keywords query with documents in vector space. But the social network is then exploited to rank the documents as more or less authoritative.

Applying to the social graph the well-known PageRank algorithm used in Google for Web pages (Brin & Page, 1998) gives a ranking for each node: highly connected individuals obtain a higher score. This score is transferred to the documents authored by each individual, and the value is then used to estimate the relevance of documents retrieved from the initial query. Documents from well-connected authors will be ranked more highly.

The approach is quite general (all kinds of social information may be exploited, as long as one can build a graph) but is clearly focused on document retrieval.

6 Conclusion

Social information retrieval is a research domain which is on the verge of maturity. A social and economic context where human resources have become the primary engine of development explains the growing interest in this recent domain. A body of techniques and applications is progressively taking form and a critical mass has now been reached.

Inspired by informal studies of human skills in locating social resources, a range of software systems has been proposed and tested in the past decade. We have at our disposal a toolbox of effective methods to retrieve human experts and socially relevant documents. Using representations and algorithms borrowed from classical information retrieval and social networks theory, social information retrieval is becoming a coherent field.

We have shown how this domain can be characterized by a generic system architecture organized around a network of personal nodes. Although variants can certainly be observed, there is now a general structure common to different research projects, structure which is strongly suggested by the specific social skills being reproduced. This common outlook, however, leaves a lot of room for various improvements and developments in the near future.

Acknowledgements: we thank F. Deloule for useful comments and suggestions.

References

- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Boston, Mass: Addison Wesley.
- Brin, S., & Page, L. (1998). *The anatomy of a large-scale hypertextual Web search engine*. Computer Science Dept., Stanford University.
- Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, Mass: Harvard University Press.
- Degenne, A., & Forsé, M. (1994). *Les Réseaux Sociaux*. Paris: Armand Colin.
- Drucker, P.F. (1992). *The Age of Discontinuity*. New York: Harper & Row.
- Goh, D., & Foo, S. (eds)(2007). *Social Information Retrieval Systems*. Hershey, PA: Idea Group.
- Kautz, H., Selman, B., & Shah, M. (1997). Referral Web: combining social networks and collaborative filtering. *Communications of the ACM* 40 (3).
- Kirsch, S., Gnasa, M., & Cremers, A. (2006). Beyond the Web: retrieval in social information spaces. *Proc. 28th European Conference on Information Retrieval*.
- Lazega, E. (1998). *Réseaux Sociaux et Structures Relationnelles*. Paris: PUF.
- Lord, M. (2007). NetSim : un logiciel de modélisation et de simulation de réseaux d'information. Mémoire de Maîtrise d'Informatique, Université du Québec à Montréal.
- Manning, C.D., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, Mass: MIT Press.
- Memmi, D., & Nérot, O (2003). Building virtual communities for information retrieval. In J. Favela & D. Decouchant (eds) *Groupware: Design, Implementation and Use*. Berlin: Springer.
- Memmi, D. (2007). The social context of knowledge. In D. Goh & S. Foo (eds) *Social Information Retrieval Systems*. Hershey, PA: Idea Group.
- Nakata, K. (2001). A grounded and participatory approach to collaborative information exploration and management. *Proc. 34th Annual Hawaii International Conference on System Sciences*. Maui, Hawaii.
- Salton, G., & McGill, M. (1983). *Introduction to Modern Information Retrieval*. New York: McGraw-Hill.
- Sun, R. (ed)(2001). Special issue on multi-disciplinary studies of multi-agent learning. *Cognitive Systems Research* 2 (1).
- Sun, R. (ed)(2006). *Cognition and Multi-agent Interaction*. Cambridge University Press.
- Voss A., Nakata, K., & Juhnke, M. (2000). Concept Indexes: sharing knowledge from documents. In M. Divitini, T. Brasethvik & D. Schwartz (eds) *Internet-Based Organizational Memory and Knowledge Management*. Hershey, PA: Idea Group.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Wooldridge, M. (2002). *An Introduction to Multi-Agent Systems*. Chichester: John Wiley and Sons.
- Zacklad, M. (2006). Documentarisation processes in documents for action (DofA): the status of annotations and associated technologies. *Computer Supported Cooperative Work* 15 (2-3).