Web Semantics for Intelligent and Dynamic Information Retrieval Illustrated Within the Mental Health Domain

Maja Hadzic, Elizabeth Chang

Digital Ecosystems and Business Intelligence Institute, Curtin University of Technology
GPO Box U1987, Perth 6845, Australia

{m.hadzic, e.change}@curtin.edu.au

Abstract. Much of the available information covering various knowledge domains is distributed across various information resources. The issues of distributed and heterogeneous information that often come without semantics, lack an underlying knowledge base or autonomy, and are dynamic in nature of are hindering efficient and effective information retrieval. Current search engines that are based on syntactic keywords are fundamental for retrieving information often leaving the search results meaningless. Increasing semantics of web content (web semantics) would solve this problem and enable the web to reach its full potential. This would allow machines to access web content, understand it, retrieve and process the data automatically rather than only display it. In this paper we illustrate how the ontology and multi-agent system technologies which underpin this vision can be effectively implemented within the mental health domain. The full realisation of this vision is still to come.

Keywords: web semantics, ontology, multi-agent systems, ontology-based multi-agent system, mental health, intelligent information retrieval, dynamic information retrieval.

1 Introduction

In 2007, the internet is frequently used as an information source for a multiplicity of knowledge domains. The general public uses Google predominately to obtain information covering these domains. The users inevitably have different access and understanding of the results they obtained from their search. As Google is not built to separate authoritative from dubious information sources, the users may have to use specialised search engines.

The accumulation of published information is an additional problem that complicates the search. For example, biomedical researchers would use PubMed which is a service of the U.S. National Library of Medicine that includes over 16 million citations from life science journals for biomedical articles back to the 1950s. Using the PubMed search engine, the user receives a list of journals related to the specified keyword. It is now left to the user to read each journal individually and try to establish links within this information. This would be easy if the journal list consisted of only a few journals. However, the journal list usually consists of thousands of

journals and the medical researchers usually do not have time to go through these results thoroughly. There is a high chance that some important information will be overlooked.

Specific and targeted searches are very difficult with current search engines. For example, a search for “genetic causes of bipolar disorder” using Google provides 960,000 hits consisting of a large assortment of well meaning general information sites with few interspersed evidence-based resources. The information provided by the government sites is not necessarily returned on the first page of a ‘google’ search. A similar search on Medline Plus retrieves all information about bipolar disorder plus information on other mental illnesses. The main problem of the current search engines is that they match specific strings of letters within the text rather than searching by meaning.

There is a need to design an intelligent search engine that would perform a search based not on keywords but on the meaning of the information. The search engine would go through the available information, understand this information and select highly relevant information as well as link this information and present it in a format meaningful to the user.

In this paper, we will briefly introduce the technologies underpinning such a search engine. We will address current issues related to information access and retrieval from the web. We will introduce the meaning of web semantics and the role of ontologies and agent technologies in the creation of a semantically rich environment. We will also illustrate this on an example from the mental health domain.

Figure 1 Information on the web covers various knowledge domains

2 Information Variety

The number of active users and information resources are increasing each day. Adding content to the web is quite an easy task and access to the web is uncomplicated and fast. The available information covers various knowledge domains and various disciplines such as astronomy, geology, biology, computer sciences, history, arts, religion and so forth (Figure 1).

Finding relevant information among all the information resources is a difficult task. The sheer volume of currently available information limits access and retrieval of required information. Current use of web content is based on browsing and syntactic key-searches.

Browsing makes use of hyperlinks to navigate from one web page to the next. In syntactic key-word searches, the key-words, made of patterns of characters, are matched to strings occurring in resources on the web. At present, these Google-like search engines are fundamental for retrieving information but the search results are often still meaningless and completely outside the domain that users are interested in.

3 Information Retrieval Hindrances

The usefulness of available data needs to be maximised. This is becoming more important each day. The organisation and management of the available information and existing information resources are the main focus.

We identified the following underlying issues with difficult data access, retrieval and knowledge sharing:

1. The increasing body of distributed and heterogeneous information;
2. There is no underlying knowledge base available to enable shared understanding of concepts to facilitate information retrieval and knowledge sharing; and
3. Databases and information resources are autonomous, heterogeneous and dynamic (content is continuously updated). There is no efficient tool to help coordinate them.

Figure 2 Information distribution

3.1 Distributed and heterogeneous information without semantics

The rapid increase in available information makes the search for needed information even more complicated. Extensive time is required to locate the appropriate information by browsing and searching. Organisations are forced to manage their knowledge more effectively and efficiently. It is becoming more and more difficult to manage and share this information amongst potential users.

The available information is distributed over various information resources. The internet contains a huge number of documents and information which covers different knowledge domains and different areas of the same knowledge domain. Lots of the available information is repeated within numerous databases. Portions of this information may be related to each other and/or may overlap. Usually, knowledge required to solve a problem is spatially distributed over different information resources as shown in Figure 2. In this situation, sharing of information becomes crucial.

The information is found in different formats that may not be compatible with each other. Two different databases may contain information on the same topic but the way this information is structured and organised within those databases may be different. In these situations, it is quite difficult to automatically compare and analyse available information. Moreover, the available data is structured in a way to suit a particular purpose of the database and associated community. Important information in one database may be found to be insignificant for another database and be removed regardless of its importance within a larger scope. The heterogeneity of the information is becoming one of the main problems in efficient information use.

3.2 Underlying knowledge base is not available

Knowledge integration methodologies are becoming more important as technology advances and interconnections between organisations grow. Efficient storage, acquisition and dissemination of knowledge require structured and standardised organisation of data. As some information needs to be reused and shared, so to do the basic concepts used to describe the information that needs to be shared. These basic concepts may be organised differently for specialised systems but the fundamental concepts must be common to all applications. A unifying framework that represents shared understanding of domain knowledge would enable intelligent management of the available information.

3.3 Autonomous, heterogeneous and dynamic information resources

Currently, information resources exist autonomously [1]. In most cases, the resources were designed for a single purpose and independently from others. Each separate knowledge domain generated its own data and, therefore, its own information sources. The integration of those autonomous information resources has been promoted during the last few years.

The information resources are quite heterogeneous in their content, data structure, organisation, information management and the like. A range of analysis tools are available but their capabilities are limited because each is typically associated with a particular database format. As each information source usually has its own structure, content and query language; available analysis tools often work on only a limited subset of data. Heterogeneity of the information resources makes their integration difficult.

One of the main issues is the dynamic nature of the internet environment. Content within the information resources is changeable as it is continuously updated and modified. The data and information can be added hourly or daily and may be removed the next minute. A solution needs to be developed that can address the issues associated with dynamic information resources.

4 Information Retrieval in Science

Information and computer sciences are becoming vital as they impact on a number of disciplines. Computers are used to store and manage the increasing amount of information. Also, computers are predominantly used for data analysis performed after (large-scale) experiments. Various sciences are experiencing those kinds of changes. As a result, the new generation of scientists are regarded as computational scientists as well.

Figure 3 Use of computers in the biomedical domain

We show an example from the biomedical domain in Figure 3. The researchers formulate specific hypotheses through the study of the contents of huge databases (arrow 1). Laboratory experiments need to be performed in order to test the hypotheses (arrow 2). This results in a collection of data. Computers are used to organise the data, perform needed calculations and present the data in a meaningful format for the users (arrow 3). Computers will also store this data and let the users access it when needed.

There is a need to create systems that can apply all available knowledge to scientific data. For any knowledge domain, the size of an existing domain knowledge base is too large for any human to assimilate. Computers are able to capture and store all this information but are unable to link this information and derive some useful knowledge from it. Better predictions can be made using the bigger knowledge base. But the predictions are still being made on a relatively small subset of the available knowledge due to the:

1. size of the available information;
2. autonomous, distributed and heterogeneous nature of the information resources; and
3. lack of tools to analyse the available information and derive useful knowledge from it.

As a result, some important information is being neglected. Sometimes, this can impact on the outcomes significantly. We need a system that can effectively and efficiently use all the available information.

Scientists would be able to progress much faster if using all the available knowledge efficiently. Usually, one particular problem can be solved by various research teams. Only when found in context with the other available information, the information provided by each

individual team shows its real value. Each research team provides a piece of information that together with the information provided by other research teams forms a complete picture.

Cooperation between different groups and professionals is also important. Different groups have different capabilities, skills and roles. Even if they all have common goals, they may be operating and executing their tasks on different knowledge levels. In most cases, the solution to a problem involves coordination of the efforts of different individuals and/or groups. There is an obvious need for these people to complement and coordinate with each other in order to provide the definition and/or solution to their common problem. Complexity of the information retrieval process is due to the:

1. huge body of information available;
2. nature and characteristics of information resources; and the
3. fact that many users still use collections of stand-alone resources to formulate and execute queries. This places a burden on the user and limits the uses that can be made of the information. Users need mechanisms to capture, store and diffuse domain-specific knowledge.

Use of internet for information retrieval is still a complex task. Most problems need to be decomposed into smaller interdependent sub-problems. A sequence of small tasks needs to be performed in order to solve the overall problem where each task may require querying of different information resources. Following the problem decomposition, the user needs to [2]:

- identify and locate appropriate information resources
- identify the content of those resources
- target components of a query to appropriate resources and in optimal order
- communicate with the information resources
- transform data between different formats
- merge results from different information resources

Figure 4 Integration of heterogeneous information

One of the biggest issues in the information integration is heterogeneity of different resources. Usually, a number of different information resources need to be queried in order to solve a problem effectively. Each information resource has its unique data structure,

organisation, management and so forth. This is represented in Figure 4 by different colours. It is difficult to integrate the information when found in different formats. Additional difficulties encountered by the users include:

a) Medical information alone totals several **petabyte** [1]. In 2007, a researcher setting a genetic human trial for 'bipolar disorder' would have needed to sift through a multitude of information from various sources (for example, Entrez Gene) to have found that loci 2p13-16 are potential positive sites for this disorder (the information originating from the research reported by Liu and his team [3]).

b) As research continues, new papers or journals are frequently published and **added** to the databases and more and more of this published information is available via the internet. Problematically, no collaborative framework currently exists to help inform researchers of the latest research and where and when it will become available.

c) Huge volumes of mental health information exist in different databases mostly having their own unique **structure**. This is equivalent to the situation in the Australian libraries before Libraries Australia (<http://librariesaustralia.nla.gov.au/apps/kss>) and the development of universal cataloguing standards. No tool is available to help manage, search, interpret, categorise and index the information in these disparate databases.

d) Most published results cover one specific topic. Usually, one must combine and analyse information regarding various topics in order to get an overall picture. Currently, no tool exists that allows examination and analysis of different factors **simultaneously**.

e) Portions of information or data on the internet may relate to each other, portions of the information may **overlap** and some may be semi-complementary. No knowledge based middleware is available to help identify these issues. 'Common knowledge', for example, may reduce the possibility of undertaking the same experiments such as examining the same region of DNA sequence by different research groups thus saving time and resources. This helps to create a cooperative environment making big research tasks coherent between different research teams. Gap analysis may also be easier.

5 Search Engines

We previously mentioned a number of factors responsible for complexity of the information retrieval process and these include:

1. the sheer volume of information and its continual increase;
2. nature and characteristics of information resources; and
3. many users still use collections of stand-alone resources to formulate and execute queries.

It is possible to design a search engine which will not be affected by the three above-mentioned problems. The sheer volume of the information and its continual increase can be handled by putting this information in a universally agreed and shared standardised format which can be 'understood' by computers and applications. The same is true for nature and characteristics of information resources. If their information content is formatted according to the universally agreed and shared format, their nature and characteristics will be transformed to support efficient and effective information retrieval. The user will not need to use a collection of stand-alone resources; all information resources together will 'act' as one big information resource that contains uniformly formatted information from various knowledge domains.

Currently, the fourth and main problem associated with complex and inefficient information retrieval is the nature and characteristics of the current search engines. Google-like search engines based on syntactic keywords are fundamental for retrieving information but the search results are often still meaningless. For example, consider the concept 'plate'. This concept has different meanings in different domains such as architecture, biology, geology, photography,

sports etc (<http://www.answers.com/topic/plate>) (Figure 5). Doing a search for 'plate' using current search engines, you may get varying results from various knowledge domains. If you try to specify a search by typing in 'plate, sports', you may still get something like: 'Everyone needs to bring a plate with food after the football on Sunday night'.

Figure 5 The term 'plate' has different meanings within different knowledge domains

We need search engines to look for word meanings within a specific context. The content of the web pages needs to be described in a way that search engines look for the meaning of the word within a specific context rather than the word itself. The content of the most current information resources is described using 'presentation-oriented' tags and the associated search engines work on this principle. In order to bring a positive transformation on this issue, the content of the information resources needs to be described using 'meaning-oriented' tags so that the search engines can be designed to search for the meaning of the information rather than simply its appearance in the text. This is one of the major concerns and goals in creating and increasing Web Semantics. We will discuss this more in the following section.

6 Web Semantics

Semantic Continuum moves from the semantically poor and easily understood by humans towards semantically rich and easily 'understood' by computer expressions. Semantics that exist only in the minds of humans who communicate and build web applications is implicit. Semantics may also be explicit and informal such as found in thesauri and dictionaries where terms are described and defined. Semantics may also be explicit and formal for humans such as in computer programs created by people. There is less ambiguity the further we move along the semantic continuum and it is more likely to have robust correctly functioning web applications.

Here, the semantics is explicit and formal for machines. This Semantic Continuum is shown in Figure 6.

Figure 6 Semantic Continuum

The current web can reach its full potential through enabling machines to be more actively involved in the web 'life-processes'. Machines need to be able to 'understand' the web content, access it, retrieve and process the data automatically rather than only displaying it. They need to be able to meaningfully collect information from various information sources, process it and exchange the results with other machines and human users (Figure 7). The web information that is semantically rich and easily 'understood' by computers can serve this purpose. Namely, the current web can reach its full potential through increasing semantics of its content; commonly referred to as Web Semantics.

Figure 7 Machines and humans exchange information

The main and the biggest project currently associated with establishing and increasing Web Semantics is the Semantic Web project. Tim Berners-Lee of the World Wide Web Consortium is leading this project. He aims to transform the current web into a semantic one where its content is described in computer-understandable and computer-processable meaning.

'The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.' (Berners-Lee et al. 2001, pp. 34-43)

'The Semantic Web is a vision: the idea of having data on the Web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications.' (www.w3c.org)

The emphasis in these definitions is on:

- well-defined meaning of the information
- machines use the information (automation)
- cooperation through data integration and reuse

The Semantic Web project aims to enable machines to access the available information in a more efficient and effective way. They are working on two different options towards the realisation of their vision. In the first option, the machine reads and processes a machine-sensible specification of the semantics of the information. In the second option, web application developers implant the domain knowledge into the software which will enable the machines to perform assigned tasks correctly.

The first two languages that played an important role in Web Semantics are: eXtensible Markup Language (XML) and the Resource Description Framework (RDF). XML allows the users to create their own tags and, in this way, enables them to add arbitrary structure to their documents. However, it says nothing about the meaning of this structure. On the contrary, meaning is expressed by RDF. The meaning is encoded in sets of three: things, their properties and their values. These sets of three can be written using XML tags and form webs of information about related things.

7 Agents for Dynamic Information Retrieval

Agents are intelligent software objects used for decision-making. They are capable of autonomous actions and are sociable, reactive and proactive in an information environment [4]. Agents can answer queries, retrieve information from servers, make decisions and communicate with systems, other agents or with users. Use of agent-based systems enables us to model, design and build complex information systems.

The Semantic Web project aims to transform the current web into a semantic one where agents can understand web content. This concept is presented in Figure 8. In this environment, an information resource will contain data as well as metadata which describe what the data are about. This allows agents to access, recognise, retrieve and process relevant information. They are also free to exchange results and to communicate with each other. In such environments, agents are able to access and process the information effectively and efficiently.

The web then becomes more than just a collection of web pages. Agents can 'understand' and 'interpret' meaning of the available information, perform complex tasks and communicate with each other and with humans. Humans can use one or several such agents simultaneously.

Figure 8 Annotation of information resources enables agents to work cooperatively

8 Ontologies for Intelligent Information Retrieval

Ontology is an enriched contextual form for representing domain knowledge. Ontology provides a shared common understanding of a domain and means to facilitate knowledge reuse by different applications, software systems and human resources [5, 6].

Berners-Lee et al. emphasise the interdependence of agent technology and ontologies [7]. Intelligent agents must be able to locate meaningful information on the web and understand the meaning of this information, to advertise their capabilities and to know the capabilities of other agents. All this can be achieved through use of ontologies; firstly, through ontology sharing between agents and secondly, through the annotation of information resources.

Agents need to communicate and interact with other agents. This requires them to share a common understanding of terms used in communication. The specification of the used terms and exact meaning of those terms is commonly referred to as the agent's ontology [8]. As this meaning must be available for other software agents, it needs to be encoded in a formal language. Agents can then use automated reasoning and accurately determine the meaning of other agents' terms.

Ontology as 'a formal and explicit specification of a conceptualisation' allows the information providers to annotate their information [7]. The annotation phase is a crucial step in creating a semantically rich environment which can be intelligently explored by users' agents. The annotation provides background knowledge and meaning of the information contained within this information resource.

9 The Relevance of Web Semantic within the Mental Health Domain

The World Health Organisation predicted that depression would be the world's leading cause of disability by 2020 [9]. The exact causes of many mental illnesses are unclear. Mental illness is a causal factor in many chronic conditions such as diabetes, hypertension, HIV/AIDS resulting in higher cost to the health system [10]. The recognition that mental illness is costly and many cases may not become chronic if treated early has lead to an increase in research in the last 20 years.

The complexity of mental illnesses adds further complications to research and makes control of the illness even more difficult. Mental illnesses do not follow Mendelian patterns but are caused by a number of genes usually interacting with various environmental factors [11]. There are many different types of severe mental illness [10], for example depression, bipolar disorder, schizophrenia. Genetic research has identified candidate loci on human chromosomes 4, 7, 8, 9, 10, 13, 14 and 17 [3]. There is some evidence that environmental factors such as stress, life-cycle events, social environment, economic conditions, climate etc. are important [10, 11].

Information regarding mental illness is dispersed over various resources and it is difficult to link this information, to share it and find specific information when needed. The information covers different mental illnesses with a huge range of results regarding different disease types, symptoms, treatments, causal factors (genetic and environmental) as well as candidate genes that could be responsible for the onset of these diseases. We need to take a systematic approach to making use of enormous amount of available information that has no value unless analysed and linked with other available information from the same mental health domain.

To overcome the currently complex and complicated situation, an intelligent and efficient information system needs to be designed that does not require researchers to sift through the same or similar results from different databases. This expert system needs to be able to intelligently retrieve information from the heterogeneous and disparate databases, and present it to the user in a meaningful way.

We propose a solution which includes the design of Generic Mental Illness Ontology (GMIO) and GMIO-based multi-agent system.

Generic Mental Illness Ontology (GMIO) can be developed to contain general mental health information. Four sub-ontologies can be designed as a part of the GMIO to represent knowledge about illness sub-groups (e.g., clinical depression, postnatal depression), illness causes (such as environmental and genetic), phenotypes (which describe illness symptoms) and possible treatments. The ontology and sub-ontologies will serve as template to generate Specific Mental Illness Ontologies (SHIO), the information specific to an illness in question (e.g., bipolar, depression, schizophrenia).

The GMIO needs to be effectively utilised within a multi-agent system. We need to define a multi-agent system architecture that will be based on GMIO and will enable the agents to collaborate effectively. A possible solution that includes different agent types (Interface, 'Manager', Information and 'Smart' agents) is represented in Figure 9. *Interface* agents assist the user in forming queries. Interface agents communicate user's request to the 'Manager' agents. 'Manager' agents decompose the overall task into smaller tasks and assign these sub-tasks to the various Information agents. *Information* agents retrieve the requested information from a wide range of biomedical databases. Each information agent may have a set of databases assigned to it. The information agents send the retrieved information to the 'Smart' agents. 'Smart' agents analyze this information, assemble it correctly and send to the Interface agent directing it back to the user as an answer to his/her query.

Within the GMIO-based multi-agent system, ontology is used at the different levels to:

- *locate and retrieve* requested information. Information content within an information resource can be described using an ontology. Only then, an agent committed to this ontology is able to 'understand' the information contained within these resources and is able to exactly locate and retrieve the information requested by a user.
- enable cooperatively working agents to *communicate* with each other during the process of the information retrieval. Use of ontology permits coherent communication between the different agents and facilitates sharing of the information among different agents.

- *analyze and manipulate* the retrieved information. In this way, the redundant and/or inconsistent information is removed. Only relevant information is selected, assembled together and presented to the user.
- *present* the retrieved information to the user in a meaningful way. The information is presented to the user in a way that makes it easier for the researcher, physician or patient to have an overview of the requested knowledge regarding human disease of interest. Moreover, the inherited organisation of ontologies adds a taxonomical context to search results, making it easier for the users to spot conceptual relationships in data.

To understand the mechanism behind these different operations we refer the interested reader to [12, 15].

Figure 9 GMIO-based Multi-agent System

As a result of user's query, the overall problem to be solved is constructed as Specific Mental Illness Ontology (SMIO) template from GMIO by Interface agents. Retrieving and adding of relevant information upon this SMIO template results in Specific Mental Illness Ontology (SMIO). This is shown in Figure 10.

Protégé [13] developed by Stanford University can be utilized for the modelling of the ontologies. JADE (Java Agent DEvelopment framework) [14] can be used to develop the multi-agent based system. JADE is a software framework that simplifies the implementation of agent applications in compliance with the FIPA (Foundation of Intelligent Physical Agents) specifications for interoperable intelligent multi-agent systems and it offers a general support for ontologies. The ontology can be stored as a computer readable description of knowledge and the agents can use this knowledge for intelligent actions.

The developed tool could be used internally in a specific organization such as a pharmaceutical company or consortium such as Universities Australia, or externally in a government-supported public information service e.g., PubMed

Figure 10 GMIO, SMIO template and SMIO

10 Conclusion

In this paper, we talked about the issues associated with current use and management of the information on the web and discussed proposed solutions to these problems. We explained the concept of web semantics, introduced the Semantic Web project and discussed its vision. We briefly introduced two major technologies that may help realise this vision: ontology and agent-based technology, and illustrated their significance within the mental health domain.

The implementation of such a synergetic system will result in positive transformation of world-wide mental health research and management to a more effective and efficient regime. This system is highly significant and innovative, and represents a new generation of information-seeking tool. The beauty of this innovative tool is that the complexity is hidden from the client. Users will simply see targeted answers to their questions. The ontology and agents are working in concert to remove complexity where it should be, not on the human client's computer screen but within the tool.

Even though the concepts of Web Semantics and the Semantic Web project have been around for awhile, the big breakthrough has not happened yet. Either we need new advanced technologies or we will find missing pieces of the puzzle in the work ahead.

References

1. C. Goble: The Grid Needs you. Enlist Now, Proceedings of the International Conference on Cooperative Object Oriented Information Systems, (2003) 589-600.

2. R. Stevens, P. Baker, S. Bechhofer, G. Ng, A. Jacoby, N.W. Paton, C.A. Goble, A. Brass: TAMBIS: Transparent Access to Multiple Bioinformatics Information Sources, *Bioinformatics*, 16/2 (2002) 184-186.
3. J. Liu, S.H. Juo, A. Dewan, A. Grunn, X. Tong, M. Brito, N. Park, J.E. Loth, K. Kanyas, B. Lerer, J. Endicott, G. Penchaszadeh, J.A. Knowles, J. Ott, T.C. Gilliam, M. Baron: Evidence for a putative bipolar disorder locus on 2p13-16 and other potential loci on 4q31, 7q34, 8q13, 9q31, 10q21-24, 13q32, 14q21 and 17q11-12', *Molecular Psychiatry*, 8/3 (2003) 333-342.
4. M. Wooldridge: *An Introduction to Multiagent Systems*, John Wiley and Sons, (2002).
5. A. Gómez-Pérez: Towards a Framework to Verify Knowledge Sharing Technology, *Expert Systems with Applications*, 11/4 (1996) 519-529.
6. A. Gómez-Pérez: Knowledge Sharing and Reuse, *The Handbook on Applied Expert Systems*, (1998) 1-36.
7. T. Berners-Lee, J. Hendler, O. Lassila: The Semantic Web, *Scientific American*, 284/5 (2001) 34-43.
8. T. R. Gruber: A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, 5/2 (1993) 199-220.
9. A. D. Lopez, C.C.J.L. Murray: The global burden of disease, 1990-2020, *Nature Medicine*, 4 (1998) 1241-1243.
10. M. Horvitz-Lennon, A.M. Kilbourne, H.A. Pincus: From silos to bridges: meeting the general health care needs of adults with severe mental illnesses, *Health Affairs*, 25/ 3 (2006) 659-669.
11. D. G. Smith, S. Ebrahim, S. Lewis, A. L. Hansell, L. J. Palmer, P. R. Burton: Genetic epidemiology and public health: hope, hype, and future prospects, *The Lancet*, 366/ 9495 (2005) 1484-1498.
12. M. Hadzic, E. Chang: Ontology-based Multi-agent systems support human disease study and control. Edited by H. Czap, R. Unland, C. Branki, H. Tianfield, in *Frontiers in Artificial Intelligence and Applications* (special issues on Self-organization and Autonomic Informatics), 135 (2005) 129-141.
13. Protégé: an Ontology and knowledge-based editor 2006, Stanford Medical Informatics, Stanford University, School of Medicine: <http://protege.stanford.edu/>
14. JADE (Java Agent DEvelopment Framework), <http://jade.tilab.com/>
15. M. Hadzic, E. Chang, P. Wongthongtham, T. Dillon: *Ontology-based Multi-agent Systems*, Springer, (to appear in 2008).