An Architecture for the Semantic Enhancement of Clinical Decision Support Systems

Eider Sanchez[1], Carlos Toro[1], Eduardo Carrasco[1], Gloria Bueno[2], Carlos Parra[3], Patricia Bonachela[3], Manuel Graña[4], and Frank Guijarro[5]

[1] Vicomtech-IK4 Research Centre, Mikeletegi Pasealekua 57, 20009 San Sebastian, Spain
{esanchez,ctoro,ecarrasco}@vicomtech.org
[2] University of Castilla-La Mancha, VISILAB group, ETSII, Spain
[3] University Hospital Virgen del Rocío, UCAi group, Spain
[4] University of the Basque Country, Computational Intelligence Group, Spain
[5] Bilbomatica, Spain

Abstract. Clinical Decision Support Systems (CDSS) are useful tools that aid physicians during different tasks such as diagnosis, treatment and patient monitoring. Multidisciplinary, heterogeneous and disperse clinical information and decision criteria have to be handled by CDSSs. For such tasks, Knowledge Engineering (KE) techniques and semantic technologies are very suitable, as they support (i) the integration of heterogeneous knowledge, (ii) the expression of rich and well-defined models for knowledge aggregation, and (iii) the application of logic reasoning for the generation of new knowledge.

In this paper we propose a generic architecture of a CDSS based on semantic technologies, which also considers the reutilization and enhancement of former CDSS in an organization. Particularly, an implementation of the proposed architecture is also presented, aiming to support the early diagnosis of AD.

Keywords: Decision support system, architecture, implementation, Alzheimer Disease.

1 Introduction

Clinical Decision Support Systems (CDSS) are active knowledge resources that use patient clinical data to generate case specific advice [1]. In other words, CDSS analyze data and present results to physicians. Such results are used for (*i*) supporting decision making during diagnosis and (*ii*) supporting treatment and patient monitoring. CDSS are massive information systems by nature while at the same time they present an arguably high complexity (in terms of computational resources) in the query construction and retrieval. Such complexity is increased when the variable values needed for decision-making are stored in disperse or heterogeneous repositories [2].

Some reported issues of CDSS are mentioned in the literature. Reported main difficulties are mainly presented while in the process of (a) integrating CDSS into clinical workflows and systems, and (b) transferring successful interventions from one system to another [3].

A. König et al. (Eds.): KES 2011, Part II, LNAI 6882, pp. 611–620, 2011.

In classical CDSS, the representation of knowledge is static, limiting the type of knowledge that can be represented [3]. Additionally, CDSS definition is specified only through explicit information enumeration (i.e. case-based systems). Hence, arguably no discovery of new knowledge is directly supported.

Another problem in CDSS is the fact that useful information for diagnosis is highly changeable. The aforementioned fact is due to the natural evolution of medical research, where new findings and advances are being continuously made. For instance, a biomarker could be rendered irrelevant, by a new discovery that supersedes it. Thus, variables and criteria of the CDSS should be often updated and for this reason, the maintainability of the system could be a critical problem, i.e. for decision support systems integrated to clinical systems [3]. Terminological interoperability is also an important matter that classical approaches in CDSS do not solve appropriately [3]. Two different CDSS may not understand each other, even if their domain and purpose is the same, because they can adopt different terminologies or, in extreme cases, due to the inertia related to monolithic and legacy system architectures.

Knowledge Engineering (KE) techniques can arguably face efficiently the aforementioned problems (which are in essence, Knowledge handling problems) because, by definition, their underlying models support (i) the integration of heterogeneous knowledge, (ii) the expression of rich and well-defined models for knowledge aggregation and (iii) the application of logic reasoning for the generation of new knowledge [4]. In particular, semantic technologies have been described in the literature as a promising approach to solve knowledge handling problems in medical domain, as shown by Gnanambal *et al.* [5] and by Yu *et al.* [6].

In this paper we propose a generic architecture for the semantic enhancement of CDSS, which also considers the reutilization of knowledge embedded in a CDSS, in order to provide the enhancement of such CDSS taking into account and lessening if not solving the main problems mentioned before. An implementation of our proposed architecture is also presented; this implementation deals with the specific domain of early diagnosis of Alzheimer Disease (AD).

This paper is arranged as follows: in section two we present briefly the related work which is relevant for our approach; in section three we introduce the architecture of a generic knowledge-based Clinical Decision Support System; in section four we present an implementation of the proposed architecture for the early detection of AD, and lastly, we present future work and conclusions in section five.

2 Related Work

In this section we present a short overview of the previous work mentioning briefly concepts related Clinical Decision Support Systems (CDSS) and the possible benefits of the application of semantic technologies in this domain.

2.1 Architectures in Clinical Decision Support Systems

According to Wright *et al.* [3] the evolution of architectures for CDSS has followed four phases: standalone CDSS, CDSS integrated to clinical systems, standards-based systems, and service models.

Standalone CDSS run separately from any other system, such as clinical systems containing the clinical information from patients and cases. Thus, a physician has to intentionally enter the required information and ask for the aid. Time is consumed during this process. Usually the system is not proactive when supporting decision making. On the bright side, these CDSS are very easy to share [3]. On the other hand, CDSS integrated into clinical systems behave just the opposite.

Standards-based systems aim to the standardization of the computerized representation, encoding, storing and sharing of clinical knowledge and decision support content [3]. There are several standards offering a different focus, such as Arden Syntax [7] and GELLO [8].

Service models separate clinical information systems and CDSS [3], and integrate them, while using standardized service-based interfaces. The standard interface can be both, located in front of the clinical system, so that any decision support system that understands the standard can make inference (i.e. HL7 vMR [9]), or located in front of the decision support system, in order that any clinical system that understands this standard can ask for aid to a known CDSS (i.e. HL7 DSS [10]).

Additionally, some other CDSS architectures have also been presented in [11],[12],[13].

2.2 Semantic Technologies Applied to Clinical Decision Support Systems

Knowledge Engineering (KE) techniques can face efficiently the aforementioned problems such as terminological interoperability, system maintainability and source heterogeneity and disparity. More precisely, semantic technologies have been described in the literature as a promising approach to solve knowledge handling in medical domain, i.e. in [5], where different approaches using semantic technologies are presented for several research directions in the medical domain.

In particular, ontologies are very promising. Gruber defined ontologies in the computer science domain as the explicit specification of a conceptualization [14]. Ontologies can fulfill efficiently the needs for organized and standardized terminologies and reusability at a structural level [15]. Because of the aforementioned fact, important consequences in the medical domain can be derived. Some results can be applied to solve interoperability issues, as shown by Ghawi *et al.* in [16], where a general interoperability architecture is presented that uses ontologies for explicit description of the semantics of information sources.

Ontologies also deliver interesting benefits, when used for reasoning and inferring new knowledge [6]. For instance, the fast query systems presented by Toro *et al.* [17].

Among the most widely used ontologies within the medical domain, we can mention the Semantic Web Application in Neuromedicine (SWAN) [18] and the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) [19].

SWAN represents an effort to provide an integrated scientific knowledge for researchers to share their results within different projects and locations. It is the result of a project intended for developing an integrated scientific knowledge infrastructure using Semantic Web technologies. SWAN has been applied to Alzheimer Disease, but it is not limited to it. The integration with SWAN endorses contents with hypotheses and publications, as shown by Lam *et al.* [18]. Decision support needs to be documented and SWAN overcomes this task.

SNOMED CT is a common standardized comprehensive clinical terminology that provides clinical content and expressivity for clinical documentation and reporting [19]. SNOMED CT provides the core general clinical terminology for the Electronic Health Record (EHR). It describes different clinical concepts such as diseases and procedures. Mapping our own ontology to a standard ontology such as SNOMED CT provides reusability of the proposed ontology, according to Houshiaryan *et al.* [15]. This fact is very important for CDSS, to overcome the lack of common language problem. Our approach shares some basic ideas of the works presented by Hussain *et al.* [11] and Lindgren [20], related to the benefits and techniques needed for the coexistence of CDSS and semantic technologies. Our main focus is the re-use and standardization of knowledge as well as the user expertise which ultimately generates the production rules that provide the diagnosis support.

3 Proposed Architecture

In this section, we propose a generic architecture of a CDSS based on semantic technologies. The proposed architecture consists of 4 layers: *Data Layer*, *Translation Layer*, *Ontology and Reasoning Layer* and *Application Layer*. This architecture considers the reutilization and enhancement of former CDSS on existence in an organization. Fig. 1 depicts an overview of our architecture.

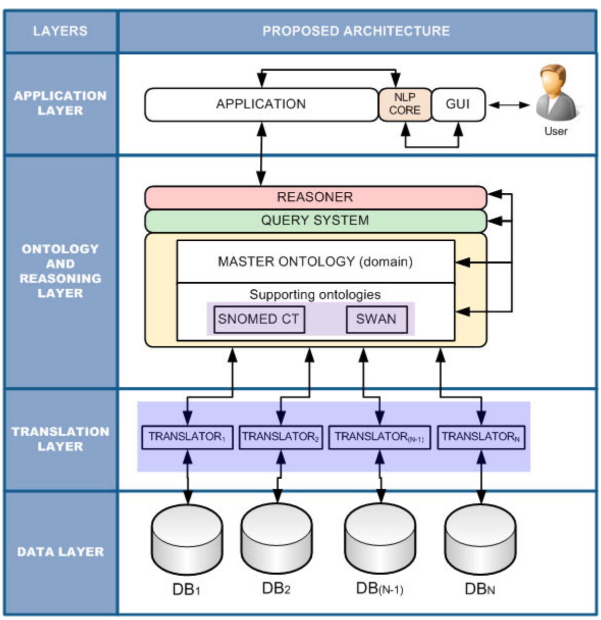

Fig. 1. Proposed architecture for the Clinical Decision Support System

3.1 Data Layer

This layer contains a collection of accessible Data Bases (DB) that store all needed information from medical systems, patient information management systems, and Picture Archiving and Communication Systems (PACS), amongst others. Is common to find that these DB are heterogeneous, as well as spatially disperse.

3.2 Translation Layer

This layer contains a translator for each DB in the *Data Layer*. Each translation module retrieves the corresponding information from the DB and aligns them into the Knowledge Bases in the *Ontology and Reasoning Layer*. In other words, the translator matches the contained information with the ontologies in the upper layer. Within this approach DB do not need to intercommunicate directly, allowing a decentralized data repository to provide input to a centralized knowledge repository.

3.3 Ontology and Reasoning Layer

The *Ontology and Reasoning Layer* deals with the knowledge embedded in the system and performs reasoning processes in order to provide diagnoses. It consists of three modules: the ontologies module, the query system and the reasoning module.

Three different ontologies giving different approaches and descriptions make up the ontologies module: a master ontology and the supporting ontologies SWAN and SNOMED CT.

SWAN links and endorses the knowledge of the system about the disease with hypotheses and publications that are being held by the medical and scientific community [18]. Therefore, the terms and the criteria in the system can be validated to be current and updated.

SNOMED CT is used to for standardization purposes [15]. This alignment provides our system with interoperability to other CDSS or knowledge sources. SNOMED CT is a general-purpose ontology and may not describe all terms needed by the system. Thus, a master ontology containing those particular terms required by each specific CDSS is proposed in our approach. This master ontology is defined by experts for the specific domain of the CDSS.

The reasoning module performs a semantic reasoning process based on expert-given rules, for the knowledge discovery. It queries the underlying ontologies through a query system. The reasoning process concludes diagnoses which are grounded and are presented to physicians to support them during decision making.

3.4 Application Layer

The interaction between the user and the system will be held by a graphical user interface (GUI). The GUI communicates with a natural language processing core, that converts to machine-language procesable the queries made by the user, and to natural language the answers returned by the system. The *Application Layer* communicates directly with the reasoner in the layer below. In this way, the output given the reasoner is presented to physicians clearly, so that a support is given in order to make decisions.

4 Architecture Implementation

The proposed architecture was implemented under the framework of the project MIND (CENIT-20081013). This project is a multidisciplinary approach to Alzheimer Disease (AD) and it is particularly focused on the early diagnosis of AD. The neurodegenerative process of AD is irreversible and for that reason a prodromal diagnosis is desirable. The common approach to support the diagnosis is based on the analysis of the results of different parameters, regarding neurological tests, neuropsychological tests, genetic studies, metabolomical studies, volumetry analysis and diagnostic image processing, amongst others [21]. During this process physicians have to deal with large amounts of heterogeneous and multidisciplinary variables. Additionally, the state of the art regarding these relevant parameters, biomarkers and procedures to follow in order to carry out a proper diagnosis is changing very fast, so that physicians have to be updated with the last medical findings.

Our architecture can handle efficiently with these tasks. Fig. 2 depicts the implementation of our presented architecture within the domain of the early detection of AD.

The *Data Layer* contains two DB: the ODEI Data Base, storing the results of the clinical tests carried out to patients, and the ODEI PACS, containing the DICOM image studies.

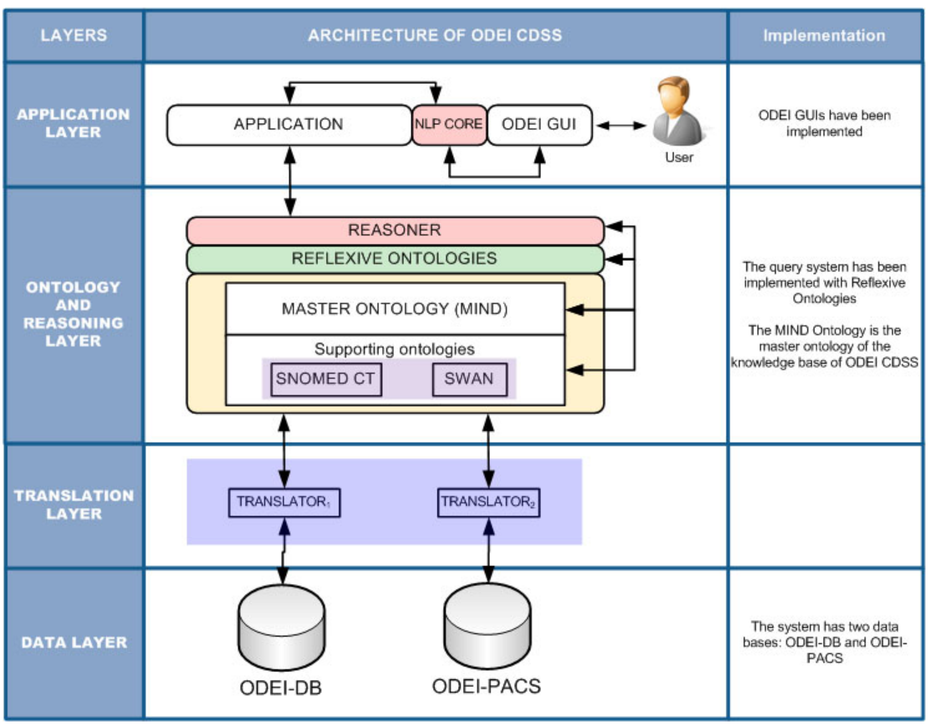

Fig. 2. Architecture implementation for the CDSS of the MIND project

The *Translation Layer* is composed by xml data models, used to align the information stored in DB with the knowledge of the upper later. For every DB an xml schema document is created containing all the elements of the DB. A data xml document is created for every data generated, at the time when data is added, modified or deleted from the DB.

The ontologies module in the *Ontology and Reasoning Layer* contains a master ontology modeling the knowledge concerning the results of the clinical tests and image studies carried out to patients. Such knowledge model is depicted in Fig. 3.

The supporting ontologies are SWAN [18] and SNOMED CT [19]. SWAN contains the description of the domain of the AD and SNOMED CT describes the patient, from a clinical point of view.

The reasoning module performs a semantic reasoning process based on rules given by domain experts, which model the process of diagnosis of AD. The query system implemented is based on Reflexive Ontologies presented by Toro *et al.* [17].

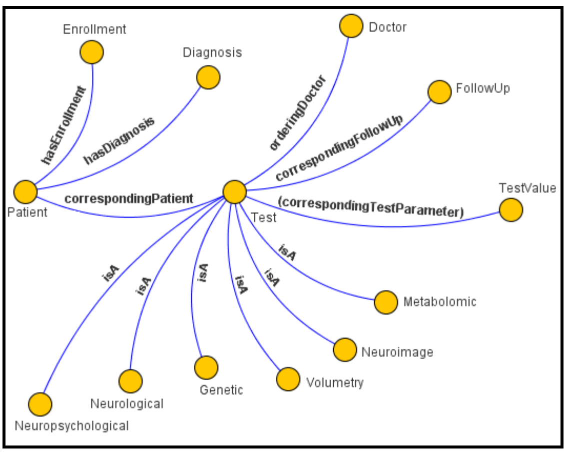

Fig. 3. Knowledge model of the results of the clinical tests and image studies

The production rules follow a classical if/then/else structure and are, on the one side, weighted depending on an importance hierarchy, and on the other, endorsed by the corresponding bibliographic source via a link given by the mapping of the MIND ontology and SWAN. Fig. 4 depicts an example of one of our production rules.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<RuleSet>
    <LoadRule>
        <Rule>if ( ( CLASS Neurological with the PROPERTY Neur
        <weight>0.6</weight>
        <AccordingTo>doi: 10.1016/S0028-3932(01)00055-0</Accor
    </LoadRule>
</RuleSet>
```

Fig. 4. Production rule example

In the *Application Layer* the ODEI Graphical User Interface (GUI) has been implemented, as well as a Natural Language Processing core, which translates the rules and their output, for the reasoner and the physician respectively. Fig. 5 shows a diagnosis for a patient supported by the ODEI system and the output of the reasoner.

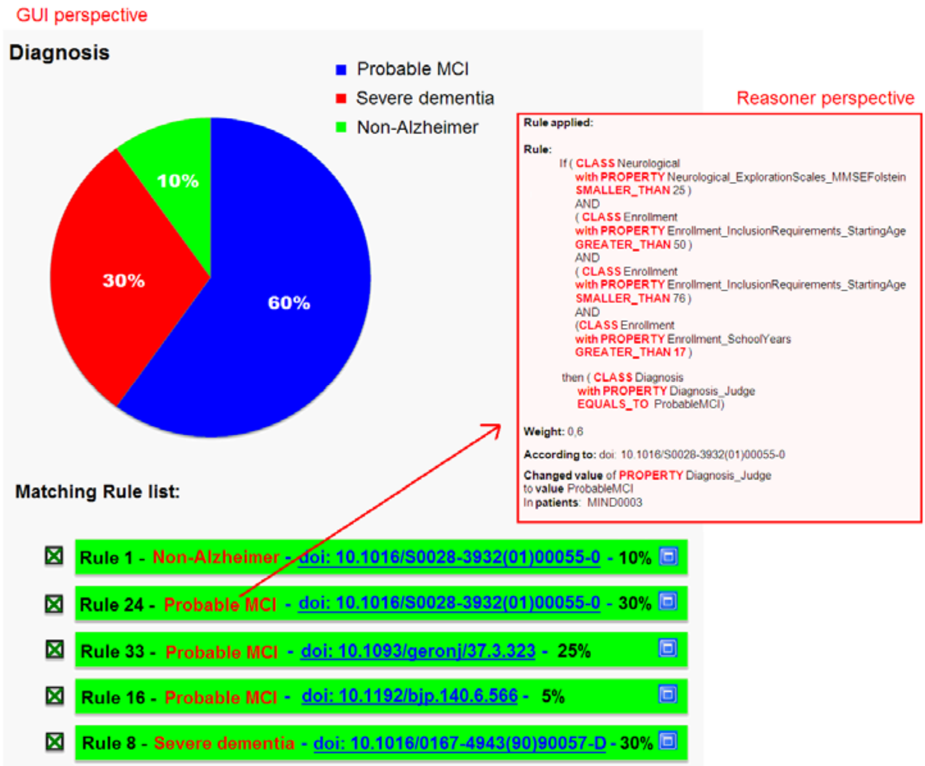

Fig. 5. Diagnosis supported by the ODEI system and output of the reasoner

5 Conclusions and Future Work

In this paper we have presented a generic architecture for the semantic enhancement of CDSS, based on semantic technologies. Our architecture allows the reutilization and enhancement of existent CDSS. In particular, this architecture has been implemented for the early diagnosis of AD. Semantics has been identified as a valuable asset technology for CDSS. The gaps bridged with the application of our methodology are related to interoperability support and discovery of new knowledge.

Our approach is horizontal and as a result the re-utilization of existent knowledge embedded in an actual CDSS is advantaged. We conclude that using standardization efforts (in the shape of existent ontologies) is beneficial as reduplication of the knowledge base is lessened and the resultant semantic layer is strengthened. In this work we used a triple ontology approach that consists of SWAN, SNOMED CT and a master ontology that contains the domain specifics that are not included directly in the aforementioned supporting ontologies.

As future work, we will explore the Set of Experience Knowledge Structure (SOEKS) [22], in order to support an experience-based reasoning that will provide an experience modeling and re-use on the production rules.

References

1. Joseph, L., Wyatt, J.C., Altman, D.G.: Decision tools in health care: focus on the problem, not the solution. BMC Medical Informatics and Decision Making 6(4) (2006)
2. Vaquero, J., Toro, C., Palenzuela, C., Azpeitia, E.: Using Semantics to Bridge the Information and Knowledge Sharing Gaps in Virtual Engineering. In: Setchi, R., Jordanov, I., Howlett, R.J., Jain, L.C. (eds.) KES 2010. LNCS, vol. 6277, pp. 495–504. Springer, Heidelberg (2010)
3. Wright, A., Sittig, D.F.: A four-phase model of the evolution of clinical decision support architectures. International Journal of Medical Informatics 77(10), 641–649 (2008)
4. Dumontier, M., Andersson, B., Batchelor, C., Denney, C., Domarew, C., Jentzsch, A., Luciano, J., Pichler, E., Prud'hommeaux, E., Whetzel, P., Bodenreider, O., Clark, T., Harland, L., Kashyap, V., Kos, P., Kozlovsky, J., McGurk, J., Ogbuji, C., Samwald, M., Schriml, L., Tonellato, P.J., Zhao, J., Stephens, S.: The Translational Medicine Ontology: Driving personalized medicine by bridging the gap from bedside to bench. In: Proceedings of the 13th Annual Bio-Ontologies Meeting, Bio-Ontologies, Boston, USA (2010)
5. Gnanambal, S., Thangaraj, M.: Research Directions in Semantic Web on Healthcare. Journal of Computer Science 1, 449–453 (2010)
6. Yu, W.D., Jonnalagadda, S.R.: Semantic Web and Mining in Healthcare. In: e-Health Networking, Applications and Services, pp. 198–201 (2006)
7. Zhou, Q., Wang, W.: The Automatic Inference of Arden Medical Logic Modules. In: Proceedings of the International Conference on BioMedical Engineering and Informatics (2008)
8. Peleg, M., Rubin, D.L.: Querying Radiology Appropriateness Criteria from a virtual Medical Record using GELLO. In: Proceedings of the Workshop on Knowledge Representation for Health-Care: Patient Data, Processes and Guidelines, in conjunction with Artificial Intelligence in Medicine Europe (2009)
9. Johnson, P.D., Tu, S.W., Musen, M.A., Purves, I.: A Virtual Medical Record for Guideline-Based Decision Support. Medinfo, 294–298 (2001)
10. Kawamoto, K., Lobach, D.F.: Proposal for Fulfilling Strategic Objectives of the U.S. Roadmap for National Action on Decision Support through a Service-oriented Architecture Leveraging HL7 Services. Journal of the American Medical Informatics Association 14(2), 146–155 (2007)
11. Hussain, S., Raza Abidi, S., Raza Abidi, S.S.: Semantic Web Framework for Knowledge-Centric Clinical Decision Support Systems. In: Bellazzi, R., Abu-Hanna, A., Hunter, J. (eds.) AIME 2007. LNCS (LNAI), vol. 4594, pp. 451–455. Springer, Heidelberg (2007)
12. Wright, A., Sittig, D.F.: SANDS: A service-oriented architecture for clinical decision support in a National Health Information Networkstar, open. Journal of Biomedical Informatics 41(6), 962–981 (2008)
13. Michalowski, W., Slowinski, R., Wilk, S., Farion, K.J., Pike, J., Rubin, S.: Design and development of a mobile system for supporting emergency triage. Methods of Information in Medicine 44(1), 14–24 (2005)

14. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies 43(5–6), 907–928 (1995)
15. Houshiaryan, k., Kim, H.S., Kim, H.H., Tung, T., Kim, I.K., Kwak, Y.S., Cho, H.: Customized Ontology-Based Intelligent Medical Terminology Tool. In: Proceedings of 7th International Workshop on Enterprise Networking and Computing in Healthcare Industry, pp. 320–324 (2005)
16. Ghawi, R., Cullot, N.: Database-to-Ontology Mapping Generation for Semantic Interoperability. In: Third International Workshop on Database Interoperability (InterDB 2007), held in conjunction with VLDB 2007, Vienna, Austria (2007)
17. Toro, C., Sanín, C., Szczerbicki, E., Posada, J.: Reflexive Ontologies: Enhancing Ontologies with self-contained queries. Cybernetics and Systems: An International Journal 39, 171–189 (2008)
18. Lam, H.Y.K., Marenco, L., Clark, T., Gao, Y., Kinoshita, J., Shepherd, G., Miller, P., Wu, E., Wong, G., Liu, N., Crasto, C., Morse, T., Stephens, S., Cheung, K.H.: Semantic Web Meets e-Neuroscience: An RDF Use Case. In: Proceedings of the ASWC International Workshop on Semantic e-Science, pp. 158–170 (2006)
19. Nyström, M., Vikström, A., Nilsson, G.H., Ahlfeldt, H., Orman, H.: Enriching a primary health care version of ICD-10 using SNOMED CT mapping. Journal of Biomedical Semantics 1(7) (2010)
20. Lindgren, H.: Towards personalized decision support in the dementia domain based on clinical practice guidelines. In: User Modeling and User-Adapted Interaction, pp. 1–30 (2011)
21. Monien, B., Apostolova, L., Bitan, G.: Early diagnostics and therapeutics for Alzheimer's disease – how early can we get there? Expert Review of Neurotherapeutics 6, 1293–1306 (2009)
22. Sanín, C., Szczerbicki, E.: Experience-based knowledge representation: SOEKS. Cybernetics and Systems 40(2), 99–122 (2009)