INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Information and Knowledge discovery system for medical data management and analysis

Marco Paulo Machado Custódio **October 2012**

Abstract

Alzheimer's disease causes cognitive decline in patients and has a devastating effect on them, causing an extremely negative socioeconomic impact in modern societies. Despite the huge amount of information available, the heterogeneity of formats and data models makes their integration and interpretation very difficult. Information Systems using Semantic Web technology allows the integration of data and knowledge spread across multiple heterogeneous sources and their use is substantially easier and more efficient, enhancing the discovery and sharing of new knowledge.

In this work, it was implemented a information and knowledge extraction system for analysis and management of medical information supported by an ontology developed for this purpose and named Neuropsychological Test Ontology. This ontology describes concepts related to the application of neuropsychological tests to patients with potentially Alzheimer's disease and was the basis for the semantic annotation of real clinical data, originally available in excel files. The developed web application performs the loading of semantically annotated data in a triple store repository allowing its search and visualization, taking into account private data criteria. In addition, interacts with a data mining software through a Web Service that makes available a service able to predict diagnoses and prognoses in a predefined time interval based on classification models, and a data import service. Through the import data service, models can be adjusted according to new data provided by the web application, reflected in the behavior of these models.

Keywords Alzheimer’s Disease, Semantic Web, Ontologies, Neuropsichological Tests, Clinical Data Integration

1. Introduction

This work, as part of the project "NEUROCLINOMICS - Understanding neurodegenerative diseases through the integration of clinical and genomic data", aimed at the design, development and implementation of an information system and knowledge extraction for analysis and management of medical information, namely information relating with neuropsychological medical tests applied to patients, potentially with Alzheimer's disease.

Due to complexity, heterogeneity and huge amount of data, scientists need new capabilities based on the modern semantic approaches. The use of Information Systems (IS) employing Semantic Web technology is currently one of the most promising and proposed solutions for data and knowledge integration, spread across multiple heterogeneous sources (Ruttenberg A., 2007). These approaches facilitate the modeling of scientific knowledge, the hypotheses checking, the development of applications that analyze heterogeneous data from different sources and domains, and sharing of new knowledge (Hey T., 2009). An interesting perspective on the handling of heterogeneous data is that making its formal definition and clarification of its meaning, the use of such data will be substantially easier and more efficient.

Thus, is still needed the development of an information system able to acquire, organize, analyze, correlate, interpret, infer and integrate heterogeneous data of neurodegenerative diseases.

2. Neuropsychological Tests Ontology

The application of neuropsychological tests aim to assess the mental health of patients, identifying and quantifying cognitive, behavioral and functional symptoms, in order to make a diagnosis, evaluate and monitor the progress of Alzheimer's disease. There is a great diversity of available tests, some of them are being applied alone or in combination with others (battery of tests). One good example is the battery of Lisbon for evaluation of dementia (BLAD) validated for the Portuguese population. The BLAD comprises tests such as *Letter Cancellation Task*, *Digit Span*, *Clock Draw*, *Interpretation of Proverbs*, *Raven Progressive Matrices, Naming Public Faces, Verbal Paired-Associate Learning, Logical Memory, Word Recall with Interference, Snodgrass and Vanderwart*.

In this work it was developed an ontology that provides the basis for the annotation of medical data concerning the neuropsychological tests. This ontology, designated by Neuropsychological Test Ontology (NTO), maps and described concepts related with

neuropsychological testing applied to patients potentially with Alzheimer's disease. This structure of knowledge representation formed the basis for the rest of the system development. Actual clinical data originally available as Excel files are semantically annotated according to NTO.

2.1. Ontology development

The concepts and their relationships taken into account for the creation of ontology are described below. A physician has a group of patients who are subjected periodically to a set of medical evaluations. Physicians and patients have a range of personal information such as, patient name, date of birth, gender, email, photography, education, etc. that could be available. The medical evaluations, performed on a specific date, corresponding to the execution of a set of neuropsychological tests. At the time of the medical evaluation, the patient has a relevant set of data, such as the duration of Alzheimer's disease, the BLAD group where they fit and their age. At the end of this evaluation, the patient will have a diagnosis relating to Alzheimer's disease, specifying the disease phase. Each neuropsychological test, may has a normalized result or not, a duration and a version. Tests may also contain various components, i.e. different parts of test, each with its duration and its own result, which may be normalized. The tests and their components may be associated with certain characteristics such as the similarity, the use of vocabulary, drawing objects, etc.. There are some more complex neuropsychological tests, called battery of tests, which group a several tests. The BLAD was the main battery of tests used in the actual data available, complemented by other tests such as the *Trail Making Test*, *Toulouse-Pierón Test* or *Californian Verbal Learning Test*.

The NTO is composed by classes (representing general domain concepts such as Patient ID, medical evaluation, etc..), Subclasses (correspond to specializations of the superclass such as Patient, Physician), instances (which represent specific data such as the phases of the Alzheimer disease, "Mild Cognitive Decline," "Moderate Cognitive Decline", etc..), properties (representing the binary relations between concepts and instances such as the property "has Medical Test" that relates Evaluation with Medical Test) and restrictions (limiting the relationships, such as, the Toulouse-Pieron test that has at maximum one test component related to dispersion index).

The class Score was created to allow the representation of data complex structure that can simultaneously have multiple dimensions such as a quantitative value, a qualitative value, a measurement unit. Thus, it is possible to represent the duration of the disease when medical

evaluation as *hasValue* = 2 and *hasUnit* = "years", or representing one of the diagnostic evaluation as *hasValue* = 1.0 and *hasQuality* = "MCI" (one of the stages of Alzheimer's disease). There are still some tests producing relevant results, not only for the value obtained in its execution, but also the time of execution. It is possible to represent these results with two instances of classes Score, the first with hasValue = 19.4, hasUnit = "in 20" and hasQuality = "Excelent" and the second with hasValue = 56 and hasUnit = "seconds".

The following figure represents the NTO conceptual model, where it can be observed the entities, their properties and the most relevant relations.

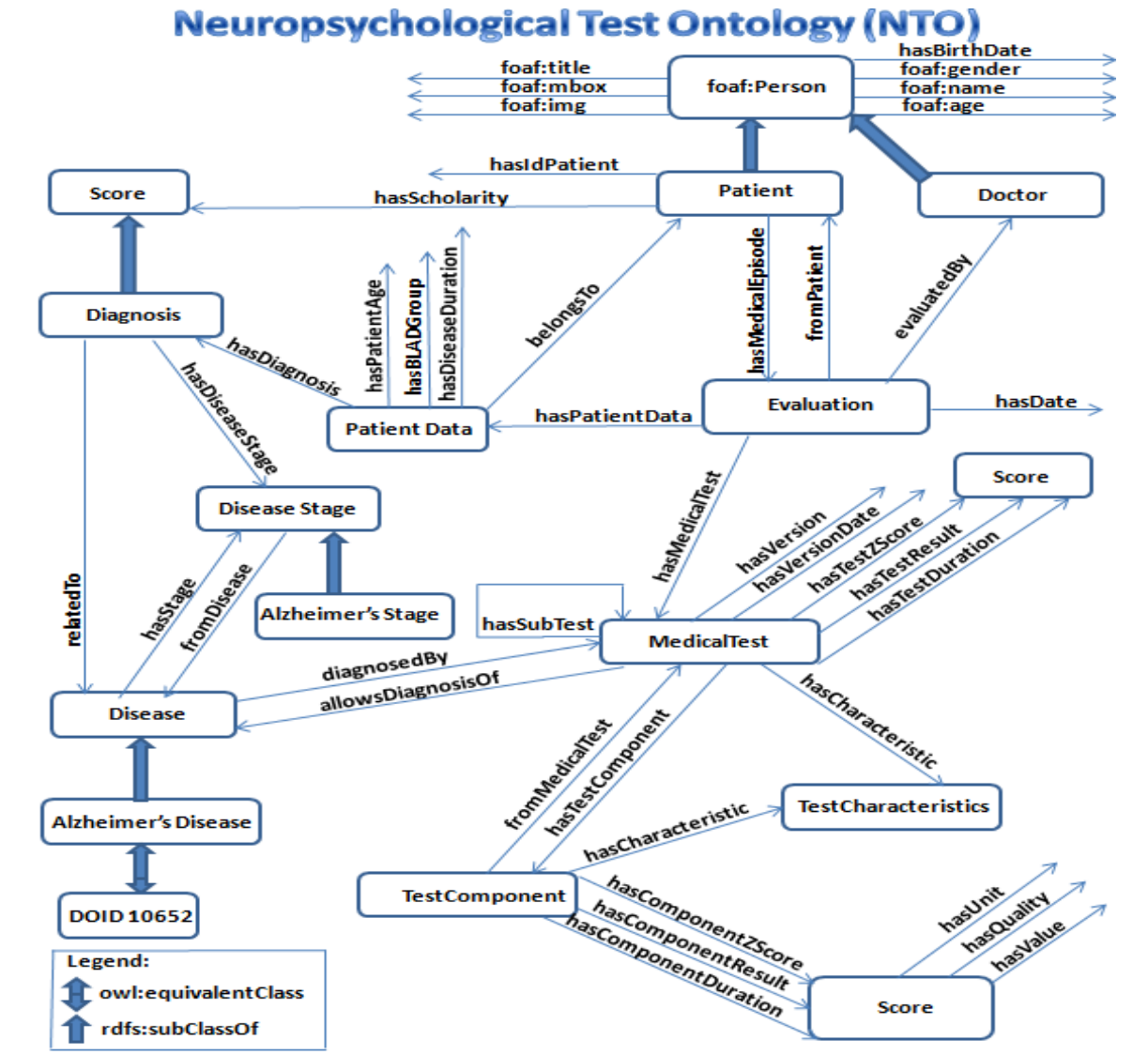

Fig. 1 – Conceptual scheme of entities and relationships of Neuropsychological Test Ontology (NTO)

3. Global vision of System

The developed prototype for the implementation of the information system and knowledge extraction, includes a Web application and a Web Service allowing users and computer systems to interact with medical data annotated according to the NTO. The system implements the Client-Server architecture with N-Tier Software Architecture, which includes 5 layers: source data layer; extract data layer; semantic data layer; semantic services layer and presentation layer. The figure below depicts the architecture.

The Web application was developed on the ZK Framework, a open source tool that allows the development of client-server applications with Ajax technology in the Java programming language. The system contains a repository of data in form of triple store implemented by *Virtuoso* server and an application server Apache Tomcat which hosted the developed Web application. The data produced by the neuropsychological tests were provided in Excel format files by physicians of the Molecular Medicine Institute (IMM). These data are then semantically annotated according to NTO and loaded into the repository. The *Virtuoso* server has a SPARQL engine and an integration module with the Jena API.

Also, a Web Service was built based on SOAP and Java EE technologies, which provides a service that relies on a data mining software, including a decision support system, based on a developed WEKA model, able to predict diagnoses and prognoses in a predefined time interval. The Web application consumes this Web Service by choosing one of the available operations. Each diagnosis and prognosis operation at 2, 3 and 4 years receives as parameters, the classifier to be used and a set of data from medical evaluation, and returns a confidence interval.

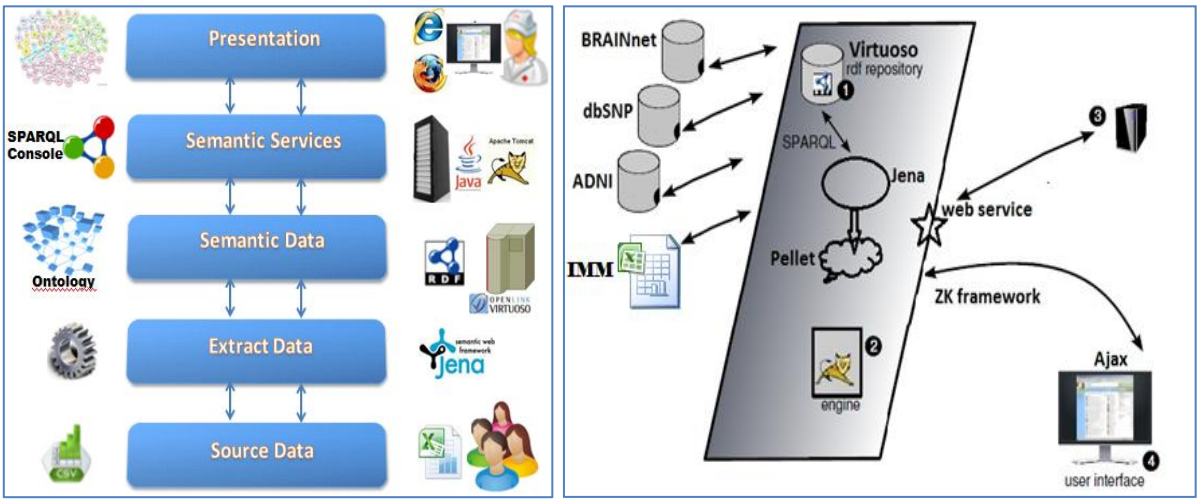

Fig. 2 –N-Tier Software Architecture and Infrastructure applicational, adapted from Kammergruber.

4. System implementation

4.1. Access control

The application uses the OpenID protocol to allow authentication of non anonymous users and Role Based Access Control that permits the application customization by the administrator in order to adapt to the needs of different user groups. In this access model, the users identified by their email, may associate to five distinct groups of users, also called for access profiles or roles: Application administrator, Physician, Secretary, KDBIO and Anonymous. Each of these roles has an associated set of permissions (read, insert, update and remove) on services provided by the application. Thus, for example, is possible to specify that only users with the Physician role can insert and remove medical records, but users with Secretary role are only allowed to update the records. The navigation options in the application, through the general menu and menu-specific, are also supported as services, allowing a personalized navigation according to their roles. Permissions are loaded into the RDF graph and stored in memory in a persistent object, which allow the rapid manipulation.

4.2. Data Annotation

This function allows the semantically annotation of data produced in neuropsychological tests, performed in patients in a particular medical evaluation, and put them in an RDF graph in *Virtuoso* system, named by DATA_GRAPH. Clinical data provided by the IMM physicians, partners in the Neuroclinomics project, are related to a set of medical evaluations conducted since 1989 to patients with potentially Alzheimer's disease. These patients underwent BLAD and a set of other individual neuropsychological tests such as *Toulouse-Pieron, Trail Making Test, Stroop Test, Wechsler Adult Intelligence Scale* and *California Verbal Learning Test*. After the evaluation step, the patients were diagnosed by the physician as "normal" or belonging to one of phases of the disease. Data provided by the IMM team in Excel files have a similar scheme to the figure below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Patient_Id	Age	DiagNPS	Diagnosis_code	Disease_duration	Date	School	Group	Gender	Birth	As_cut	As_time	As_tot	DS_forw	DS_back	DS_tot
4	2	64	MCI	1,0	1,00	28-11-2002	16	7	1	26-01-1938				5	4	10
5	2	65	MCI	1,0	1,50	05-06-2003	16	8	1	26-01-1938				5	5	12
6	2	66	MCI	1,0	3,00	22-12-2004	16	8	1	26-01-1938	16	37	4,3	5	4	11
7	2	68	MCI	1,0	5,00	04-01-2007	16	8	1	26-01-1938	16	34	4,7	5	4	11
8	2	71	MCI amnésico	1,0	7,50	15-04-2009	16	8	1	26-01-1938	14	40	3,5	5	3	9
9	3	68	MCI md	1,0	3,00	16-04-2001	4	5	1	13-09-1932	15	55	2,7	5	2	7
10	4	83	MCI amnésico	1,0	4,00	09-01-2008	16	8	1	09-10-1924	15	34	4,4	5	4	9
11	5	69	MCI	1,0	5,00	06-12-2006	4	5	1	17-09-1937	16	39	4,1	5	3	8

Fig. 3 – Excel file with medical data

To allow flexibility to the system, allowing the exchange of columns, changing column names, and the addition of more columns to accommodate new tests, it was created a configuration mapping ontology file, designated by OntologyMap.xls. This file allows semantic annotations in each column according to the ontology NTO.

	A	B	C	D	E	F	G	H
1	*File Column Name / Ontology*	M_Calc	Ident	As_cut	As_time	As_tot	WAIS_cubos	WAIS_semelhanças
2	*Component Test Name*			AS_Cut	AS_Time	AS_Total	WAIS_Cubes_Component	WAIS_Similarities_Component
3	*Test Name*	MentalCalculationTest	ObjectIdentificationTest	LetterCancellationTask	LetterCancellationTask	LetterCancellationTask	WAIS-III	WAIS-III
4	*Score Property Name*				hasUnit			
5	*Score Property Value (seconds, etc)*				seconds			
6	*Zscore Header Name*		Identification_Z			CancellationTask_Z	CubesWAIS_Z	
7	*Data Type of Value (if different from decimal)*							String

Fig. 4 –Configuration OntologyMap.xls file

4.3. Patient management

This function allows the management of general information concerning patients. In this screen, data such as name, date of birth, photograph, nickname, title, email and phone can be inserted or updated in the system. Patients are also associated with their physician.

The confidentiality of personal data and medical tests associated with real patients is a subject of great criticism. This problem was solved by performing a separation in terms of graphs between personal data and data from medical tests. From one side, the medical tests data are stored in the graph DATA_GRAPH and correlated with a number that belongs to a patient ID. Using this graph, it is impossible to identify the real patient. From the other side, the personal data of patients are kept in the graph FOAF_GRAPH and will not be accessed through SPARQL console. This graph also contains the patient ID, enabling the association between these data and the data from graph DATA_GRAPH.

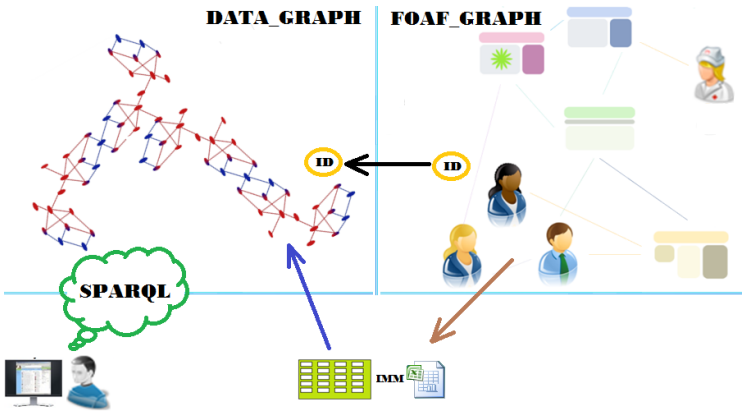

Fig. 5 – Separation between clinical data and personal data graphs in *Virtuoso*

4.4. Data Visualization

The function allows the visualization of the recorded data from medical tests. The user can choose a set of filters as the patient ID, exact date of the medical evaluation, medical testing, patient gender, etc. This feature, when performed by a user with physician profile allows to view patients information such as patient ID and his name. Thus, the physician can select the patient whose data want to view, avoiding the problem of confidentiality of the data in relation to their own patients.

4.5. Diagnosis and Prognosis

This function allows the interaction of this application with decision support system for able of predicting diagnosis and prognosis in a predefined time interval. The Mild Cognitive Impairment (MCI) is considered an early stage of Alzheimer's disease. Patients diagnosed with MCI has an increased likelihood of progression to the later stages of the disease. The support system for the decision is a data mining software, based on a developed model WEKA, which processes a wide range of real data concerning the neuropsychological tests of patients already diagnosed with MCI or AD, applying them classification models created by techniques of data mining. The import of new data provided by the web application through Web Service allows these models to be rearranged changing their behavior.

For a particular medical evaluation, the system allows to assign automatically the patient's diagnosis regarding the phase of Alzheimer's disease. In cases where the patient is still in the stage of Mild Cognitive Impairment (MCI) is even possible to make a prognosis at 2, 3 and 4 years of disease progression in patients. This feature needs to consume a Web Service that provides diagnostic and prognostic operations. Each of these operations takes as parameters the classifier to be used and a set of data produced in a clinical evaluation. The available classifiers are Neural Networks (NN), Naive Bayes (NB), K-Nearest Neighbor (KNN), Support Vector Machines (SVM) Kernel Polynomial (Poly), SVM Gaussian Radial Basis Function (RBF) and Decision Tree J48.

Each operation of the Web Service returns a confidence interval between 0 and 1. For the diagnosis, the model returns 0 means the patient has 0% of probability of being in the phase MCI, and if the first return means that the patient has a 100% of probability of being in this phase. In prognosis at 2, 3 or 4 years, if the model return 1 means that the patient has 100% of probability to not progress in the disease while 0 means it has 0% probability of not progressing in disease, this is, it will evolve in the disease.

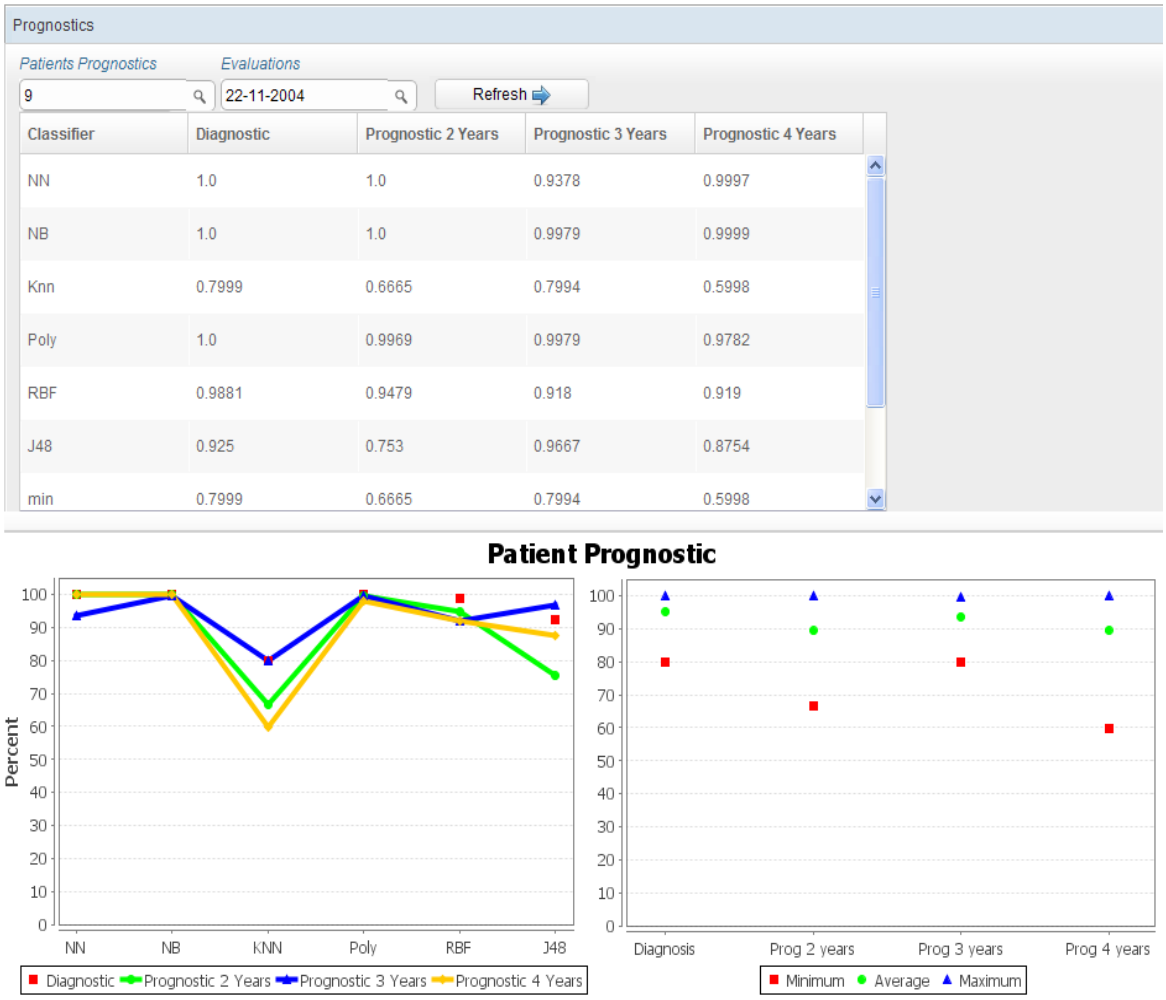

Fig. 6 –Prognostic function

4.6. Browsing in Ontology

This functionality allows browsing through Neuropsychological Test Ontology displaying its hierarchy of classes, instances, properties, annotations and data types.

4.7. Sparql queries

The user can query the graph DATA_GRAPH that resides on the server Virtuoso. This graph contains the NTO ontology and RDF data from the annotated medical tests. The user has an useful and flexible tool, that it can interact with the semantic data and obtain the desired responses.

4.8. Neuropsychological tests

This feature allows an intuitive view of psychological tests on existing ontology NTO, its hierarchy and features such as purpose, description, comments, recommended age of patients, and the test duration. It is still possible to observe a pdf file with the chosen test description.

5. Conclusions

In this work it was implemented a prototype of an information system and knowledge extraction for analysis and management of medical information. The *Neuropsychological Test Ontology* was developed allowing the description and relation of concepts on application of neuropsychological tests to patients with potentially Alzheimer's disease. This structure of knowledge representation was the basis for the semantic annotation of real data, originally available as Excel files provided by IMM medical team.

The web application contains a set of features available including, the introduction of new patients and physicians in the system; loading new data into the annotated triple store repository; viewing information regarding medical evaluations; querying data through a SPARQL interface, interaction with a Web service that allows to assign automatically the diagnostic to the patients and make predictions about the evolution of Alzheimer's disease.

The import of data provided by the web application through the Web Service allows adjustment of the models created by techniques of data mining.

5.1. Contributions and future work

The created ontology represents the effort to formulate a rigorous and comprehensive conceptual scheme within the area of medical neuropsychological tests knowledge. When there is an agreement in respect of semantic concepts and their relations, is possible to share data and knowledge in a coherent and consistent across various applications of this domain. Thus, it is open the way to annotate the data from neuropsychological tests performed by patients around the world, using a simple and user-friendly system, allowing that other physicians and researchers can easily search and extract new knowledge related with neuropsychological tests.

As future work, propose the inclusion of clinical data from external sources (BrainNet, ADNI and dbSNP) in the repository. These data should be semantically annotated according to the NTO in order to make available a set of integrated data, allowing their extraction and sharing of new knowledge. Furthermore, the ontology developed was mainly focused on neuropsychological tests performed in patients with potentially Alzheimer's disease. The extent of this ontology to other neurodegenerative diseases such as the case of Amyotrophic Lateral Sclerosis disease will be very useful and could be done in the future.

References

Adamusiak, T., Burdett, T. e Kurbatova, N. 2011. *OntoCAT - simple ontology search and integration in Java, R and REST/Javascript.* s.l. : BMC Bioinformatics, 2011.

Antezana, E., et al. 2009a. *BioGateway: A Semantic Systems Biology Tool forthe Life Sciences.* s.l. : BMC Bioinformatics, 2009.

Antezana, E., et al. 2009b. *The Cell Cycle Ontology: an application ontology for the representation and integrated analysis of the cell cycle process.* s.l. : Genome Biology, 10:R58, 2009/2.

Antezana, Erick, Kuiper, M. e Mironov, V. 2009. *Biological knowledge management:the emerging role of the semantic web technologies.* s.l. : Briefings in Bioinformatics, 10(4):392–407, 2009.

Aranguren, M. E., et al. 2008. *Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology.* s.l. : BMC bioinformatics, 9(Suppl 5):S1., 2008.

Belleau, F., et al. 2008. *Bio2RDF: Towards a mashup to build bioinformatics knowledge systems.* s.l. : Journal of Biomedical Informatics, 2008.

Berners-Lee, Tim, Fielding, R. e Masinter, L. 2005. *Uniform resource identifier(uri): Generic syntax.* s.l. : RFC3986, 2005.

Berners-Lee, Tim, Hendler, J. and Lassila, O. 2001. *The Semantic.* s.l. : Scientific American, 284, 34-43., 2001.

Bishop, Matt. 2005. *Introduction to Computer Security.* s.l. : Addison-Wesley, 2005. ISBN: 0-321-24744-2.

Bizer, C., Heath, T. e Berners-Lee, T. 2009. *Linked Data - The Story So Far.* s.l. : International Journal on Semantic Web and Information Systems, 2009.

Bizer, Heath, Tom e Christian. 2011. *Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136.* s.l. : Morgan & Claypool., 2011.

Boisot, M. e Canals, A. 2004. *Data, information and knowledge: have we got it right?* s.l. : Journal of Evolutionary Economics., 2004. Vols. 14(1):43–67.

Bono, B., et al. 2011. *The RICORDO approach to semantic interoperability for biomedical data and models: strategy, standards and solutions.* s.l. : BMC Bioinformatics, 2011.

Cali, Andrea, et al. 2001. *Accessing data integration systems through conceptual schemas.* s.l. : Lectures Notes in Computer Science, 2001.

Chaves, Ana Paula e Steinmacher, Igor. 2011. *OntoDiSENv1: an Ontology to Support Global Software Development.* Campo Mourão, Brasil : Federal Technological University of Paraná, 2011.

Cheun, Kei-Hoi, et al. 2007. *Semantic Web Approach to database integration in the life sciences.* s.l. : in C. J. O. baker & Kei-Hoi Cheun Semantic Web, Springer US, pp. 11-30, 2007.

Clercq, Jan De. 2002. *Single Sign-on Architectures.* s.l. : in Infrastructure Security International, 2002.

Clermont, Gilles., et al. 2009. *Bridging the gap between systems biology and medicine.* 2009. 1(9):88.1-88.6.

Erling, Orri e Mikhailov, Ivan. 2007. *RDF support in the Virtuoso DBMS.* s.l. : In Conf. on Social Semantic Web., 2007.

Ferraiolo, David F. e Kuhn, D. Richard. 1992. *Role-based access controls.* s.l. : In Proceedings of the 15th National Computer Security Conference, 1992.

Gordon, E. 2007. *Integrating genomics and neuromarkers for the era of brain-related personalized medicine.* s.l. : Personalized Medicine., 2007. 4(2):201-215.

Gruber, T.R. 2007. *Ontology. Disponível em: <http://tomgruber.org/writing/ontology-definition-2007.htm>.* s.l. : Acedido em 17SET2012, 2007.

Guarino, Nicola. 1998. *Formal ontology and information systems.* Italia : International Conference on Formal Ontology, 1998.

Hey, T., Tansley, S. e Tolle, K. 2009. *The Fourth Paradigm: Data-Intensive Scientific Discovery.* 2009.

Hunter, P., et al. 2010. *A vision and strategy for the virtual physiological human in 2010 and beyond.* s.l. : The Royal Society, 2010.

Kammergruber, W. C. e Ehms, K. 2010. *A Corporate Tagging Framework as Integration Service for Knowledge Workers.* s.l. : Proceedings of I-KNOW, 2010. pag. 4.

Kienast, R. e Baumgartner, C. 2011. *Semantic Data Integration on Biomedical Data Using Semantic Web Technologies.* s.l. : Bioinformatics - Trends and Methodologies, Mahmood A. Mahdavi (Ed.), ISBN: 978-953-307-282-1, InTech, 2011.

Kitano, H. 2002. *Systems Biology: a brief overview.* s.l. : Science, 2002. 295:1662-4..

Molle, R. e Haarslev, V. 2003. *Racer: An OWL Reasoning Agent for the Semantic Web.* Halifax, Canada : In Proceedings of the International Workshop on Applications, Products and Services of Web-based Support Systems, in conjunction with the 2003 IEEE/WIC International Conference on Web Intelligence, 2003.

Noorbakhsh, F., Overall, C. M. e Power, C. 2009. *Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology. Trends in Neurosciences.* 2009. 32(2):88-100.

Noy, N. e McGuinness, D.L. 2001. *Ontology development 101: A guide to creating your first ontology. Technical Report KSL-01-05.* Stanford. : Stanford Medical Informatics, 2001. DOI:10.1.1.136.5085.

Noy, N. F., et al. 2009. *BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res.BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res.* 2009.

Ouksel, Aris M. e Sheth, Amit. 1999. *Semantic Interoperability in Global Information Systems.* s.l. : SIGMOID Record 28(1):5-12, 1999.

OWL Web Ontology Language Reference. **W3C. 2004.** s.l. : W3C recommendation, 2004. disponível em http://www.w3.org/TR/owl-ref/.

Recordon, David e Reed, Drummond. 2006. *OpenID 2.0: A Platform for User-Centric Identity Management.* Alexandria, Verginia, USA, : in Conference on Computer and Communications Security Proceedings of the second ACM workshop on Digital identity management, 2006. ISBN: 1-59593-547-9..

Ruttenberg, A., et al. 2007. *Advancing translational research with the Semantic Web.* s.l. : BMC Bioinformatics, 2007.

Sandhu, R. S., et al. 1996. *Role-based access control models.* s.l. : IEEE Computer, 29(2)., 1996.

Schadt, E. 2009. *Molecular networks as sensors and drivers of common human diseases.* s.l. : Nature 461, 218-22, 2009. Nature, 461:218-223, February 2009.

Signore, Oreste. 2006. *Representing Knowledge in the Semantic Web.* 2006. W3C Italy.

Sirin, E., et al. 2007. *Pellet: A practical OWL-DL reasoner.* s.l. : Jornal of Web Semantics, 5(2):51–53, 2007.

Smith, B., et al. 2007. *The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration.* s.l. : Nature Biotechnology 25, 1251 - 1255., 2007.

Staab, S. e Studer, R. 2009. *Handbook on Ontologies.* s.l. : International Handbooks on Information Systems, second ed., 2009.

Strauss, Esther, Sherman, Elisabeth M. S. e Spreen, Otfried. 2006. *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary.* Oxford University Press : Third Edition, 2006.

Tsarkov, D. e Horrocks, I. 2006. *FaCT++ Description Logic Reasoner: System Description.* s.l. : Lecture Notes in Computer Science. 4130. pp. 292–297, 2006.

Wache, H., et al. 2001. *Ontology-Based Integration of Information — A Survey of Existing Approaches.* Seattle, WA : in Proc. of the IJCAI-01 Workshop: Ontologies and Information Sharing, 2001. pp. pp. 108-117.

Wongthongtham, Pornpi, Chang, Elizabeth e Dillon, Tharam. 2007. *Ontology modelling notations for software engineering knowledge representation.* s.l. : In: IEEE International Conference on Digital Ecosystems and Technologies, 2007.