Faceted Ontological Model for Brain Tumour Study

Subhashis Das* and Sayon Roy**

*University of Trento, ICT Doctoral School, Via Sommarive, 9 I-38123 Povo,
Trento Italy TRENTO 38123, <subhashis.das@unitn.it>

**Indian Statistical Institute, Documentation Research and Training Centre,
Bangalore India, <sayon@drtc.isibang.ac.in>

Subhashis Das was born in Kolkata, India. He obtained his master's degree in library and information science from the Documentation Research and Training Centre, Indian Statistical Institute, Bangalore, India in 2013. Currently, he is a PhD student at ICT (Information and Communication Technology)-Doctoral School, University of Trento, Italy.

Sayon Roy was born in Kolkata, India. He obtained his master's degree in library and information science from the Documentation Research and Training Centre, Indian Statistical Institute, Bangalore, India in 2014. Currently, he is a library trainee at British Council Library, Kolkata, India.

Das, Subhashis and Roy, Sayon. "Faceted Ontological Model for Brain Tumour Study." *Knowledge Organization* 43 no. 1: 3-12. 18 references.

Abstract: The purpose of this work is to develop an ontology-based framework for developing an information retrieval system to cater to specific queries of users. For creating such an ontology, information was obtained from a wide range of information sources involved with brain tumour study and research. The information thus obtained was compiled and analysed to provide a standard, reliable and relevant information base to aid our proposed system. Facet-based methodology has been used for ontology formalization for quite some time. Ontology formalization involves different steps such as identification of the terminology, analysis, synthesis, standardization and ordering. A vast majority of the ontologies being developed nowadays lack flexibility. This becomes a formidable constraint when it comes to interoperability. We found that a facet-based method provides a distinct guideline for the development of a robust and flexible model concerning the domain of brain tumours. Our attempt has been to bridge library and information science and computer science, which itself



involved an experimental approach. It was discovered that a faceted approach is really enduring, as it helps in the achievement of properties like navigation, exploration and faceted browsing. Computer-based brain tumour ontology supports the work of researchers towards gathering information on brain tumour research and allows users across the world to intelligently access new scientific information quickly and efficiently.

Received: 19 August 2015; Revised 30 November 2015; Accepted 10 December 2015

Keywords: brain tumour, ontology, faceted

1.0 Introduction

To date, ontologies pertaining to various domains and diverse dimensions have been constructed by a large number of researchers, scholars and developers across the globe. The twenty-first century witnessed a boom in the field of ontology construction. During the initial phase, ontologies were constructed for disciplines like biology

and computer science, however ontologies gradually began developing for other subjects and domains.

According to the World Health Organization (WHO) estimates, brain tumours annually claim 400,000 victims. MRIs and computed tomography (CT) scans are mainly used for detecting and locating brain tumours, which is vital for surgery and treatment. The scope of our work is to develop an ontological model, which would provide

information on not only brain tumours and their types but also on specialists for diagnosis and treatment. We have used the methodology and guiding principles originally proposed by Giunchiglia et al. (2012).

In this paper, we have demonstrated ontology-based modelling on brain tumour research using a faceted approach, a well-known method in library and information science, which is generally used for classifying different domains. We have also presented an information retrieval system which is quite formidable when it comes to lightweight ontologies. However, our system precisely provides right information related to brain tumours, starting from brain tumour types to available treatments, doctors, etc. Our objectives:

- Depict how to build relations between different concepts and aspects of brain tumours and retrieve meaningful information to cater various information needs.
- Achieve results which would strengthen ontology based information retrieval and facilitate semantic query answering.

The paper has been organized into five sections. The introductory section provides an initial discussion on certain aspects of the paper like the purpose, scope, methodology and system design: description of Brain Tumour Ontology (BTO). It also covers the search techniques within the knowledge base; discussions have also been provided regarding BTO integration with ontology browser, which technically in information science is known as a faceted web-based browser. The concluding section presents the results achieved during the different stages of ontology construction, along with provisions for future work.

2.0 Literature review

Several studies have been carried out regarding the detection of brain tumour, and different ideas have been put forth by various researchers based on their opinions and methodologies used for its detection so that a diversified inference can be obtained regarding this disease and its causes and remedies. Various techniques are being applied at present, which are based on soft computing (Logeswari and Karnan 2010). Other work, such as a high-speed parallel fuzzy logic C-mean algorithm for brain tumour segmentation has been carried out by Murugavalli and Rajamani (2006). Murugavallil and Rajamani (2007) have come up with an improved implementation of brain tumour detection technique using segmentation based on neuro fuzzy technique.

In recent times brain tumour ontology has gained popularity, thus Sinha and Mathur (2012) improved brain

tumour detection with principles of ontology. Their work was mainly focussed on the detection of brain tumours with the aid of digital image processing from MRI images. They also followed nine steps (including thresholding, watershed segmentation, morphological operation) for detecting brain tumours, matching them with their existing database of brain tumour images. They used Protégé (http://protege.stanford.edu/) software for creating classes along with their attributes. The main purpose of their research was to analyse the given image with the one already present in their knowledge database, so that the inference could be made as accurate as possible. Hermi'T tool was used for analysing images.

Diagnosis using MRI and MRS is the main way to detect brain tumours. MRI imaging in India is a costly business. However, if equipped with accurate and precise information then prescribing MRI for detection of tumour existence might be avoided. Our system will make an effort to answer with the maximum possible accuracy to amateurs and medical practitioners who are doubtful about the existence of brain tumours for any particular patient. Hu et al. (2011) presented their experiences of representing the knowledge behind Health Agents, a distributed support system for diagnosis of brain tumours based on the ontology developed by them. For building the HA domain ontology (HaDOM), they used the web ontology language (OWL) for its reasoning support and expressiveness. HaDOM captured expertise information for facilitating diagnosis and prognosis of different types of brain tumours and other issues related to tumour patients. Conceptual structure was prepared after discussion with different domain experts such as neuro-surgeons, biochemists and oncologists. This allowed representation of knowledge from different domains as well as patient management during the proper course of treatment. It is very easy to map all the attributes of a particular patient with onto-graphical user interfaces. HaDOM has core concepts from the domain and the top-level conceptual relations between those concepts. haMedCrtl.owl, ha-Classifier.owl, haSecurity.owl, CNS_Anatomoic_Structure .owl are the four modules of HaDOM. HaDOM provides a common interface among different hospitals, which maintain their own database. RDFQ (RDF Data Query Language) is used for retrieving information from a database. This paper presents one declarative framework to separate the functionality of the system from an articulate interface with respective derivations from user requirements and/or queries.

In the doctoral thesis of Khotanlou (2008), a new method for segmentation of pathological brain structures and image information (region and edge) for segmentation was proposed. For representing prior information, they used the ontological engineering tool Protégé and

captured the knowledge of brain tumours in three distinct ontologies viz.—brain anatomy ontology, brain tumour ontology and special relation ontology. They also proposed a simple ontology for a specific classification of tumours. For example, Swanson et al. (2003) and Clatz et al. (2005) have recently introduced bio-mathematical models to quantify and describe the growth rate of gliomas, visualized radiologically. The model in Swanson et al. takes into account the two major biological phenomena underlying the growth of gliomas at the cellular scale: proliferation and migration.

Hadzic and Change (2005) explained generic human disease ontology, and our work had its initial inspiration from that piece of research. Our ontology that has been developed would primarily be of use to physicians and researchers; however, its primary aim is to assist beginning neurosurgeons. By entering the symptoms into the system, surgeons may be able to retrieve the information respective to a particular disease. Furthermore, the system also can help those physicians identify the disease and consider available treatments. Our system is ably equipped to display all possible treatment options available for that particular disease.

Several studies (e.g., Lobach and Hammond 1997) have demonstrated that a medical practitioner shows more effective compliance with guidelines when they are embedded in the knowledge layer of the clinical decision support system (CDSS) and provide a customized userfriendly interface for an individual patient at the point of care. For helping non-specialist health practitioners (Lobach and Hammond 1997), we have clinical practice guidelines (CPG), where disease-specific recommendations are made, which assist clinical decision-making in accordance with best evidence. In rural areas and in the interior parts of the country where a specialist health practitioner is rarely available, a CPG based decision support system is the best guide for medical practitioners. Clinical practice guidelines for providing clinical decision support of breast cancer by Abidi (2007) presented a semantic web approach to develop a clinical decision support system to support family physicians who are basically nonspecialist to provide breast cancer follow-up care. They used Protégé ontology editing environment to build their breast cancer ontology by making use of OWL. In their paper they classified the concept of Breast cancer into eight classes (e.g. Patient_Type, Physician_Type, etc.) and demonstrated the potential of semantic web technology for developing a CPG-driven decision support system. Chen, Bau and Huang (2010) discussed technological developments in medical science using OWL and Semantic Web Rule Language (SWRL) for knowledge representation about the diagnosis and treatment of diseases.

3.0 Purpose

The purpose of this work is the development of an ontology-based framework for building an information retrieval system which would create specific queries for users. For this ontology creation, information has been obtained from a wide range of information sources in the brain tumour literature with subsequent compilations of numerous individual aspects. This compiled information has been analysed to provide a standard, reliable and relevant information base for our proposed system.

4.0 Methodology

In this section, we describe the main steps and guiding principles, which are based upon those originally propounded by Giunchiglia, et al. (2012). Before creating an ontology on any domain, one needs to possess a clear and vivid idea about what the ontology can address:

- Who are the brain tumour specialists and what are their areas of expertise?
- Which doctors have specialized knowledge on brain tumours and in which hospitals are they available?
- What are the different hospitals in the country specialized in dealing with brain tumours?
- What are the different symptoms?
- Who are the brain tumour specialists in Kolkata?
- Which varieties of brain tumours have rapid growth and become fatal overtime?
- Which brain tumours are most common among men and women in their 60s-80s?
- Which categories of brain tumours are extremely common among children?

4.1 Problems of current information retrieval systems

The major problem with current information retrieval systems is high recall with low precision as recall and precision have an inverse relationship. Different search engines are mainly based on keyword search mechanisms. The results thus retrieved display all the documents wherever that particular term occurs. Hence, many junk and irrelevant documents are retrieved. For example, someone seeking information on which hospitals in Kolkata specialize in treatment of brain tumours will get millions of records but few are useful among them. The solution for such a problem is to develop an ontology-based information retrieval system, along with a corresponding search mechanism.

4.2 Steps of building ontology

4.2.1 Step 1 identification of terminology

For our work, technical terms have been collected from various sources published by different brain tumour associations and/or societies. As main sources of natural language terminology, we have selected "about brain tumour: a primer for patients and caregivers" by the American Brain Tumour Association®. Terms have also been taken from a classification by The American Association of Neurological Surgeons (AANS). Then, they have been grouped based on similar attributes or characteristics. Each formal term is then classified as a class, sub-class, individual, relation or attribute.

4.2.2.Step 2 analysis

The formal terms collected during the previous step were analysed for identification of their commonalities and differences. They were also analysed to identify those terms that will serve as a class, property (object or data), or instance. Then those terms were used as building blocks for construction of a brain tumour ontology. A literature review and browsing different web pages also helped in identifying different standard terms or preferred terms. For example:

- Gliomas, Pilocytic astrocytoma, Glioblastoma multiforme, Optic nerve glioma
- Specialist, Doctor
- Vomiting, Double vision,
- {Contract no, concept Id, ICSD no}
- CT scan, MRI Scan

4.2.3 Step 3 synthesis

Synthesis involves arrangement of facets according to their similarity in characteristics and labelling the categories. After divison into classes, they were arranged in hierarchical order (super-class, sub-class).

- Brain tumour types
 - Primary brain tumour
 - Secondary brain tumour
- Brain tumour cause
 - Environmental cause
 - Genetic cause

4.2.4 Step 4 standardization

A standard (or preferred) term should be selected among the terms collected from the literature. The terms chosen should most frequently be used or collected from standard or controlled vocabularies available from that domain. For our work, we have standardized terms from SNOMED CT®. In some instances, we have used terms which were most frequently used rather than those from controlled vocabularies.

4.2.4 Step 5 ordering

The ordering of terms should be based on the purpose, scope and subject of the ontology. We have ordered (or rather, arranged) the brain tumour class according to the AANS classification scheme. A limitation is that for our work, we have only standardized terms, which were used to build the brain tumour type class apart from SNOMED CT ®.

4.3 Brain tumour ontology using the concept DERA

DERA (Domain, Entity, Relation and Attribute) is a faceted knowledge organization framework. It makes a provision for organization of knowledge into facets by defining them as per their domains (Giunchiglia and Dutta 2011; Giunchiglia et al. 2014). In DERA, domains consist of three elements, namely entity (E), relation (R) and attribute (A). D=<E, R, A>. Now we would like to describe the brain tumour ontology from the DERA perspective. In this ontology, brain tumour is a domain (D), which contains a class, relation between classes or objects and attribute or characteristic for refining class or entity.

Entity, as stated by Giunchiglia and Dutta (2011) is an elementary component that consists of classes (categories) and their instances, having either perceptual correlates or only conceptual existence in a domain in context. This entity definition is slightly different from Bhattacharyya's (1975) definition of entity although the main idea is derived from it. An example of entity class is:

Gliomas

- Lowest grade tumours
- Lower grade malignancies
- Higher-grade malignancies
- Highest-grade malignancies
 - Glioblastoma multiforme
 - Gliosarcoma
 - Gliomatosis cerebri

Relations(R) build connections between two classes or entities. Let us consider an example:

Retrieve all the doctors who are specialists in brain tumours and in which hospital they are available. Knowl. Org. 43(2016)No.1 7



Figure 1. Object property in BTO



Figure 2. Data property in BTO

In our BTO ontology, relation "isAvailableIn" connects two classes: "Brain_tumour_Specialist" and "Hospital." In Protégé, this type of relation is called an object property (Figure 1).

Attribute (A) is mainly the characteristic of the classes. Data type attribute qualifies or quantifies the entity (Figure 2). For example, "has_SNOMED_CT®_ConceptId and characterizes" provides more attributes about brain tumours.

4.4 System design using Protégé 4.2

Figure 3 explains the overall process from identification of terms to search query formulation along with standardization and ontology building.

4.4.1 Composition of BTO

For ontology construction, we used Protégé 4.2. We have built the ontology using all the features (such as structural hierarchy), properties and other features (such as restrictions) made available in the ontology editor. We have defined seven main classes: namely, Brain_tumour_Type,

Brain_tumour_Cause, Brain_tumour_Diagnosis, Brain_tumour_Institution, Brain_tumour_Symptom, Brain_tumour_Treatment, and Brain_tumour_Specialist. The seven main classes have further been divided into subclasses as the need arose. In the cases above, a few classes belong to two superclasses. For example, "Pineal_Cyst" is a subclass of two classes: "Cyst" and "Pineal_Tumour." "Rathke's_Cleft_Cyst" is a subclass of "Cysts" and "Pitutary_Tumour," and "Pineoblastoma" is a subclass of both "Pineal_Tumour" and "Primitive_neuroectodermal_tumors (PNET)." "Gliosarcoma" is a common member of "Higher-grade malignancies" and "skull_base_primary_tumour." The class hierarchy for the brain tumour ontology is shown in Figure 4.

4.4.2 Specifying properties for the BTO classes

The Brain_tumour_Type is the most important class, because based on different locations of tumours, respective treatments are suggested. For specifying attributes and relations among individuals or classes, we used two types of properties available in Protégé, namely object and data properties. Object property links two individuals together.

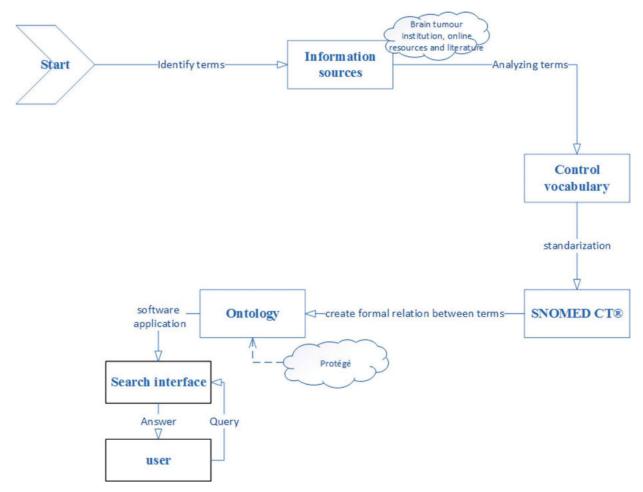


Figure 3. System design

For example, the object property "treatedBy" links an individual "Choroid_plexus_papilloma" (an instance of the class "Braintumour_Type") with another individual "Removal_of_the_tumour_by_Surgery" (an instance of the class "Brain_tumour_Treatment"). The data properties link individuals to data type or value. For example, an individual "Glioblastoma multiforme" can have attributes (e.g., "has_ICDO_3_code," "has_SNOMED_CT®_conceptId," etc.) and values (e.g., "9440,9441,9442/3" and "63634009").

4.4.3 Specifying property characteristics

Certain properties such as "has_ICDO_3_code," "has_SNOMED_CT®_conceptId" were deemed as functional properties since a term (for example, different brain tumours) can have only one ICDO (International Classification of Diseases for Oncology) code, or/its? SNOMED_CT® identity number.

4.4.4 Individuals of BTO ontology

Following the next step of BTO ontology engineering development, we specified the individual based on collected

information. We modelled relationships among the classes using the object properties, which connect two classes or individuals. Data type properties allowed the embodiment of character of an individual. Inter-class relationships between different classes are shown in Figure 5.

Here, the relationship is shown as follows: "Dr. Amitabha Candra" is an individual of class "Brain_tumour_specialist," with the relation "isAvailableIn," and it is connected with "Fortis_Hospital_Kolkata," an individual of the Hospital class.

4.4.5 Problems and solutions during working with Protégé 4.2

By clicking on OWLVIZ for visualization, class and subclasses appear, and dot.error may be found. The solution to this problem is to first install Graphiviz 2.28 (Window). Then, in the Protégé 4.2 File (tab) preferences, OWLVIZ, a dot application, add the absolute path where this "dot.exe" file is located, e.g. in case Windows OS (C: \programFile\Graphviz2.28 \bin\dot.exe).

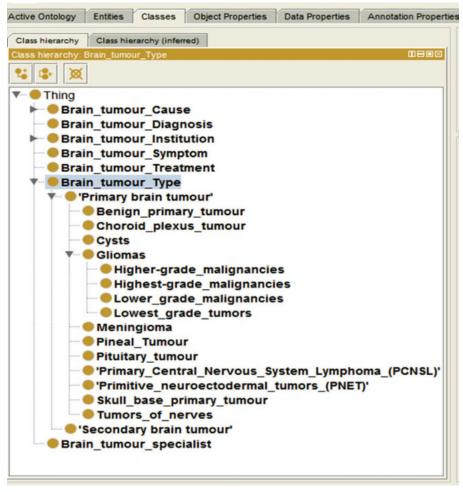


Figure 4. BTO class hierarchy

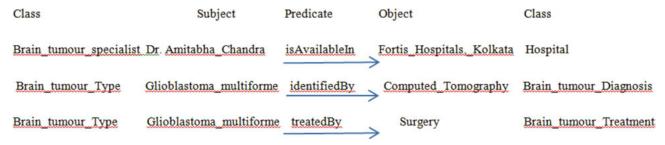


Figure 5. Inter-class relationship in BTO

5.0 Ontology queries and visualization

In this section, we have elaborated how the BTO can be used as a knowledge base to answer a query. Let us consider one simple query:

SELECT ?DoctorName ?HospitalName WHERE {?DoctorName http://www.semantic web.org/samsung/ontologies/2012/10/untitledontology-9#isAvailableIn> ?HospitalName } Which are the brain tumour specialist Hospitals in India? The query is expressed using SPARQL language as follows:

9

PREFIX bt:SELECT ?Hospital_name ?Country_from WHERE { ?Hospital_name bt:Country ?Country_from .

FILTER (?Country_from = "India"^^xsd:string). }

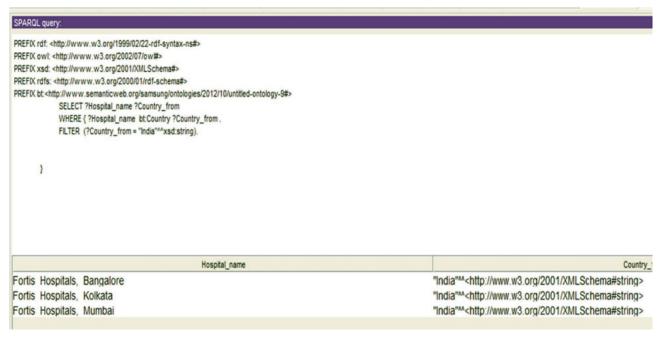


Figure 6. Query execution

The query is executed using the HermiT reasoner, and the result is as show in Figure 6.

6.0 Integration of BTO with ontology browser

A search interface should have all the features that can guide a user in finding the required information. It should be easy to use. Writing an unambiguous query is a difficult task to achieve. Also, it is not expected that a general user can formulate SPARQL or any other query language.

We have used Ontology Browser v1.4.4 (http://owl.cs. manchester.ac.uk/repository/browser), a web based ontology browser created as part of the CO-ODE project by the University of Manchester. The important features of the Ontology Browser v1.4.4 are:

- It loads content from anyplace on the web.
- It enables us to browse and visualize any complex OWL ontology.
- It is quite easy to integrate the OWL ontology with Ontology Browser.
- It has many tabs, namely ontologies, classes, object properties, data properties, annotation properties and individuals for navigating your ontology as per the requirement.
- Fully indexed and linked content are split down into module and entity type.
- To find entities by name with autocomplete feature.

For indexing an ontology, one has to keep the ontology in one cloud storage-based ftp site. Figure 7 depicts all in-

formation about Dr. Deepu Banerji, such as his expertise in surgery, his first name (Deepu) and his availabilty at Fortis Hospital Kolkata.

7.0 Conclusion

Representing active knowledge about brain tumours is quite important and largely advantageous. A computer based brain tumour ontology supports the work of researchers in gathering information on brain tumour research and allows users across the world to intelligently access new scientific information quickly and efficiently. Shared knowledge improves research efficiency and effectiveness, because it helps to avoid unnecessary redundancy in research, thereby avoiding repetition of work. Our ontology will facilitate the exact combination of the genetic and environmental factors involved as well as their individual influence on brain tumours.

We wish to build an ontology for one particular disease whose mortality rate is very high and collect information or data from the best hospitals that provide specialized service. Our acumen lies in the creation of an information retrieval system (which should be able to retrieve any answer) and which simultaneously would also work as a semantic search engine atop the developed ontology. This ontology will help to guide all new medical practitioners as well as laymen who seek information. Our system will provide the best results with high precision.

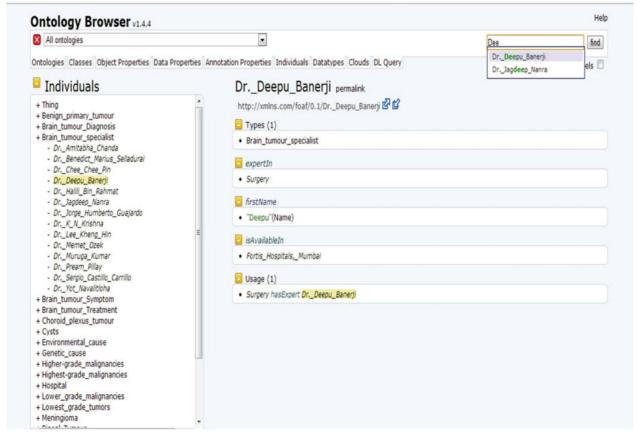


Figure 7. BTO ontology visualization

Notes

1. URL for accessing braintumour.owl file http://web protege.stanford.edu/#Edit:projectId=eb0bc51a-539b-42e5-8e13-535c0a333cfe

References

Abidi, Samina Raza. 2007. "Ontology-based Modeling of Breast Cancer Follow-Up Clinical Practice Guideline for Providing Clinical Decision Support." In *Proceedings of the 20th* IEEE International Symposium on Computer-Based Medical Systems, June 20-22 2007. Washington, D.C.: IEEE, 2007, 542-47.

Bhattacharyya, Ganesh. 1975. "Fundamentals of Subject Indexing Languages." In Ordering systems for global information networks: Proceedings of the Third International Study Conference on Classification Research, January 6-11 1975, Bombay, India, edited by A. Neelameghan. [The Hague]: FID/CR, 83-99.

Chen, Rung-Ching, Cho-Tsan Bau and Yun-Hou Huang. 2010. "Development of Anti-Diabetic Drugs Ontology For Guideline-Based Clinical Drugs Recommend System Using OWL and SWRL." In *Proceedings of the 2010 IEEE International Conference on Fuzzy Systems*

(FUZZ), July 18-23 2010, Barcelona, Spain, Piscataway, N.J.: IEEE, 1-6.

Clatz, Olivier, Maxime Sermesant, Pierre-Yves Bondiau, Herveé Delingette, Simon K. Warfield, Grégoire Malandain and Nicholas Ayache. 2005. "Realistic Simulation of the 3-D Growth of Brain Tumors in MR Images Coupling Diffusion with Biomechanical Deformation." IEEE Transactions on Medical Imaging 24: 1334-46.

Giunchiglia, Fausto and Biswanath Dutta. 2011. "DERA: A Faceted Knowledge Organization Framework." Submitted to the International Conference on Theory and Practice of Digital Libraries 2011 (TPDL'2011) http://eprints.biblio. unitn.it/2104/1/techRep457.pdf

Giunchiglia, Fausto, Biswanath Dutta, Vincenzo Maltese, and Feroz Farazi. 2012. "A Facet-Based Methodology for the Construction of a Large-Scale Geospatial Ontology." *Journal on Data Semantics* 1: 57-73.

Giunchiglia, Fausto, Biswanath Dutta and Vincenzo Maltese. 2014. "From Knowledge Organization to Knowledge Representation." Knowledge Organization 41: 44-56.

Hadzic, Maja and Elizabeth Chang. 2005. "Ontology-based Support for Human Disease Study." In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences, January 3-6 2005,* edited by Ralph H Sprague. Los Alamitos, Calif.: IEEE Computer Society Press, 143a.

- Hu, Bo, Madalina Croitoru, Roman Roset, David Dupplaw, Miguel Lurgi, Srinandan Dasmahapatra, Paul Lewis, Juan Martinez-Miranda and Carlos Saez. 2011. "The HealthAgents Ontology: Knowledge Representation in a Distributed Decision Support System for Brain Tumours." The Knowledge Engineering Review 26: 303-28.
- Khotanlou, Hassan. 2008. "3D Brain Tumors and Internal Brain Structures Segmentation in MR Images." PhD diss., Télécom ParisTech.
- Lobach, David F. and W. Ed Hammond. 1997. "Computerized Decision Support Based on a Clinical Practice Guideline Improves Compliance with Care Standards." *The American Journal of Medicine* 102: 89-98.
- Logeswari, T. and M. Karnan. 2010. "An Enhanced Implementation of Brain Tumor Detection Using Segmentation based on Soft Computing." In *Proceeding of 2010 International Conference on Signal Acquisition and Processing, 9-10 February 2010, Bangalore, India.* Los Alamitos, Calif.: IEEE Computer Society, 243-7.

- Murugavalli, S. and V. Rajamani. 2006. "A High Speed Parallel Fuzzy C-Mean Algorithm for Brain Tumor Segmentation." *BIME Journal* 6, no. 1: 29-33.
- Murugavalli, S., and V. Rajamani. 2007. "An Improved Implementation of Brain Tumor Detection Using Segmentation Based on Neuro Fuzzy Technique 1." *Journal of Computer Science* 3: 841-6.
- Ontology Browser. Retrieved June 1, 2014, from http://owl.cs.manchester.ac.uk/repository/browser
- Protégé. Retrieved February 11, 2015, from http://protege.stanford.edu/
- Sinha, Monika, and Khushboo Mathur. 2012. "Improved Brain Tumor Detection with Ontology." *International Journal of Computational Engineering Research* 2: 584-8.
- Swanson, Kristin R., Carly Bridge, J. D. Murray, and Ellsworth C. Alvord. 2003. "Virtual and Real Brain Tumors: Using Mathematical Modeling to Quantify Glioma Growth and Invasion." Journal of the Neurological Sciences 216: 1-10.