The African Journal of Information Systems

Volume 10 | Issue 3 Article 3

May 2018

Semantic Data Storage in Information Systems

Jean Vincent FONOU DOMBEU
University of KwaZulu-Natal, fonoudombeu@gmail.com

Raoul KWUIMI
Vaal University of Technology, kwuimi@gmail.com

Semantic Data Storage in Information Systems

Research Paper

Volume 10, Issue 3, July 2018, ISSN 1936-0282

Jean Vincent Fonou Dombeu
University of KwaZulu-Natal
fonoudombeu@gmail.com

Raoul Kwuimi
Vaal University of Technology
kwuimi@gmail.com

(*Received September 2017, accepted February 2018*)

ABSTRACT

The storage and retrieval of information are important functions of information systems (IS). These IS functions have been realized for decades, due to the maturity of the relational database technology. In recent years, the concept of Semantic Information System (SIS) has emerged as IS in which information is represented with explicit semantic based on its meaning rather than its syntax to enable its automatic and intelligent processing by computers. At present, there is a shortage of discussions on the topic of semantic data storage in IS as compared to the relational database storage counterpart. This study uses a combination of qualitative and quantitative methods to discuss semantic data storage in IS. The qualitative method is by means of literature review to learn the existing techniques for representing and storing semantic data. The quantitative method is done with experiments to empirically discuss these techniques. The empirical findings of the study shed light on the technologies and approaches utilised to store semantic data in relational databases. This may contribute to the understanding of semantic technologies in IS and foster the development of semantic information systems.

Keywords

Information Systems, Semantic Information Systems, Ontology, Semantic Data Storage

INTRODUCTION

An information system is defined by Fill (2009a) as an ensemble of interdependent components that assist organizations in decision making and control, through the collection or retrieval, processing, storage and distribution of information. This definition shows that the storage and retrieval of information are important functions of information systems. These IS functions have been achieved efficiently in the previous decades due to the maturity of the relational database technology. In fact, the relational database technology has existed for more than 30 years and enjoys widespread adoption through a multitude of books prescribed at academic institutions in various courses. The Relational Database Management Systems (RDBMS) provides sophisticated functionalities for data storage management and retrieval such as query optimization, transaction processing, access control and data recovery for organisations and industries (Heymans, Ma, Anicic, Ma, Steinmetz, Pan, Mei, Fokoue, Kalyanpur, Kershenbaum, Schonberg, Srinivas, Feier, Hench, Westzstein, & Keller., 2008; Wilkes, Hoover, Keer, Mehra, & Veitch., 2006).

The relational database technology prescribes the Entity Relationship Diagram (ERD) to capture the business logic of the IS and to represent the data that is to be stored. The ERD is further converted into relational database schemas that are used to create the database and store the IS data through RDBMS. Although the database schemas represent the relationships between the data based on the business requirements of the IS domain captured in the ERD, they do not represent any explicit semantic of the data. Furthermore, in today's competitive business market, companies do not have to use information within their local IS but they also need to share and use information from the information systems of other companies (Guido & Paiano, 2010). The semantic representation of IS data would enable computers to interpret and automatically process these data, thereby, enabling the integration and interoperability between different information systems of organizations. Moreover, the semantic representation of IS data would enable computers to automatically reason these data and infer new knowledge from them to support decision making in the organisations.

In recent years, the concept of Semantic Information System (SIS) has emerged as IS in which information is represented with explicit semantic based on its meaning rather than its syntax to enable its automatic and intelligent processing by computers (Fill, 2009b). In the SIS, the meaning of information is represented with ontology. An ontology is a simple view of a domain through its concepts, entities and objects, and the relationships between them; it provides a common representation of knowledge or data in a domain to facilitate information sharing and integration amongst heterogeneous information systems. The ontology of an IS domain needs to be further represented formally in a language that can be understood and processed by computers. The two most popular of these languages are Resource Description Framework (RDF) and Web Ontology Language (OWL). Further discussions of these languages are provided later in this study.

In summary, the data in the SIS is represented with ontology. Ontology is referred to as semantic data, due to its formal representation in RDF/OWL for computer processing and reasoning (Fernandez, Arias, Martinez-Prieto, & Gutierrez, 2013). Therefore, hereinafter, the terms ontology and semantic data are used interchangeably. Furthermore, the task of building ontology is out of the scope of this paper; some methodologies for building ontologies that capture the semantic relationships between the data of the IS domains can be found in Pinto, Staab and Tempich (2004), Sure, Staab and Studer (2002) and Suarez-Figueroa (2010). Instead, this study uses ontologies downloaded from the internet to empirically discuss the existing techniques for representing and storing semantic data.

Over the past few years, different techniques have been developed to store semantic data in computer memory, in file systems and relational databases (Lili, Lee, & Kim, 2010; Ramanujam, Gupta, Khan, Seida, & Thuraisingham, 2009). However, there is a shortage of discussions of these techniques for storing semantic data in IS as compared to the relational database storage techniques; this may undermine the development of SIS. This study discusses the topic of semantic data storage in IS. A combination of qualitative and quantitative methods is used. The qualitative method is by means of a literature review to learn about the existing techniques for representing and storing semantic data. The quantitative method is done with the design research method to conceptualize, design and conduct experiments to empirically discuss these techniques. The empirical findings of the study shed light on the technologies and approaches utilised to store semantic data in relational databases. This may contribute to the understanding of semantic technologies in IS and foster the development of semantic information systems.

The rest of the paper is organized as follows. Section 2 discusses the related studies. The methodology of the study is explained in Section 3. The existing technologies and approaches for storing semantic data are presented in Section 4. Section 5 describes the experimental results of the study and a conclusion ends the paper in the last section.

RELATED WORK

The topic of semantic technologies adoption in IS has been of interest to many researchers in recent years (Fill, 2009b; Guido & Paiano, 2010; Nalepa & Furmanska, 2009; Rajbhandari, Gosai, Shah, & Pramod, 2012; Ricca, Grasso, Liritano, Dimasi, Lelpa, Manna, & Leone, 2010; Soualah-Alila, Faucher, Bertrand, Coustaty, & Doucet, 2015; Yadagiri & Ramesh, 2013).

Guido and Paiano (2010) proposed the use of ontology to integrate information systems of various domains. A shared global ontology is built to represent the data in the information system domains; the global ontology provides a shared and common representation of the semantic of the data in these information systems, thereby, enabling one information system to access information from another without any prior agreement. This idea was extended in Medical Information Systems (MIS) by Rajbhandari et al. (2012) who proposed a solution for the integrated access to patients' information in heterogeneous MIS of hospitals based on semantic technologies. The authors proposed the use of ontology to provide a shared and common representation of patients' information and the storage of the resulting ontology in a central server where clients and medical doctors can seamlessly access patients' records remotely from any hospital.

Solutions for adopting semantic technologies in the tourism domain are also addressed (Ricca et al., 2010; Soualah-Alila et al., 2015). Ricca et al. (2010) developed a tourism ontology to model the process of organizing and selling holiday packages to clients. The tourism ontology represents data related to the geographic, travel agent knowledge, user preferences, and tourism offer information. The resulting tourism ontology was further implemented in a logic based programming language; this enabled reasoning of the ontology to select suitable holiday packages for customers. An ontology-based solution is proposed by Soualah-Alila et al. (2015) to facilitate the task of finding and publishing tourism data on the Web. The concepts and relations for the tourist resources are modelled with an ontology, namely, TIFSem. The TIFSem ontology was further implemented and queried to provide answers to tourist requests.

Other experiences of using semantic technologies in Geography Information Systems (GIS) and the library are presented in Nalepa and Furmanska (2009) and Yadagiri and Ramesh (2013), respectively. Nalepa and Furmanska (2009) present existing metadata and ontologies of the GIS domain and explained how they can be integrated to provide a distributed and collaborative environment for secure access to GIS data. Semantic technologies are discussed in Yadagiri and Ramesh (2013) and they show how it can be utilised to improve services and access to the Library collection.

The concept of Semantic Information System (SIS) is the focus of research in Fill (2009b). The author defined SIS as an IS in which information is pre-encoded with semantics, enabling both humans and machines to interact and process it. The author went on to propose an approach for modelling SIS; the approach consists of extending an e-business modelling framework, namely, E-BPMS by integrating ontologies and semantic description of social interaction of users and machines.

Although some of the related studies discussed above (Guido & Paiano, 2010; Rajbhandari et al., 2012) have mentioned the storage of ontology, none of these studies has addressed the technologies and approaches required to store the resulting ontologies in IS. This research overcomes this shortcoming of the previous work and conducts an empirical discussion of existing technologies and approaches for storing semantic data in IS. The methodology of the study is presented in the next section.

METHODOLOGY

A combination of qualitative and quantitative methods is used in this study. The qualitative method is by means of a literature review to identify relevant publications pertaining to existing approaches for storing and querying semantic data as well as the platforms and storage media for implementing these approaches. Thereafter, the quantitative research is carried out with experiments using the Design research method.

The design research method consists of a set of predefined steps that enables the solving of a problem or the creation of new knowledge (Hevner & Chatterjee, 2015). These steps consist of five activities namely: awareness, suggestion, development, evaluation, and conclusion (Vaishnavi & Kuechler, 2004). In this study, the awareness stage identifies the need for a framework to empirically discuss semantic data storage in IS. In the suggestion stage, the framework is designed and specified. The required platforms are identified to implement the framework in the development stage. The evaluation stage analyses the results of the implementation of the framework and the conclusion stage discusses the underlying approaches and structures used to store semantic data.

Figure 1 presents the framework developed in the suggestion phase of the Design research method to empirically analyse the semantic data storage approaches.

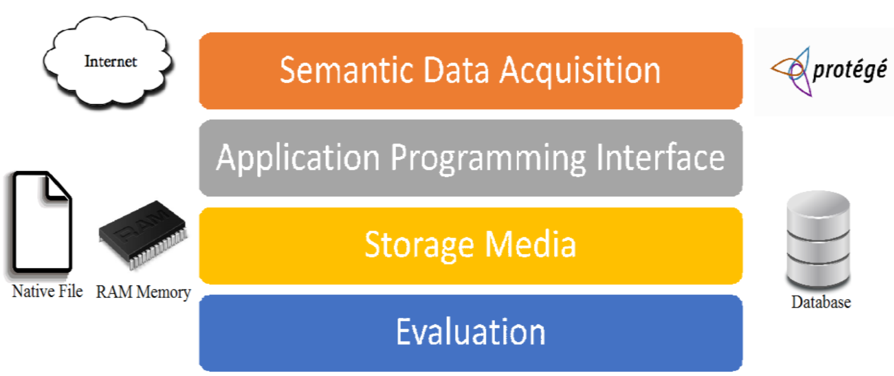


Figure 1. Framework for Semantic Data/Ontology Storage

The first layer of the framework is the Semantic Data Acquisition layer. The purpose of this layer is to acquire the ontologies or semantic data that will be used by the other layers. As mentioned earlier, the development of ontology is out of the scope of this research. The Semantic Data Acquisition layer of the Framework in Figure 1 acquires existing ontologies that have been developed and made available publicly on the internet. The second layer is the Application Programming Interface (API) layer; it is used to create, edit, browse and delete ontology or semantic data. It is also used to load existing semantic data. The third layer which is the Storage Media, is used to physically store the semantic data in the computer memory. Evaluation is the last layer; it analyses and discusses the underlying structures used to store semantic data.

SEMANTIC DATA STORAGE

This section presents the literature related to semantic data storage in terms of the languages used to encode the semantic data, the approaches that are utilised to store semantic data and the existing software platforms for storing and querying semantic data.

Languages for Representing Semantic Data

In today's competitive business environments, most IS data of organisations is accessible by customers via the internet. However, information in the current internet is only interpreted and understood by human beings; this makes a large amount of information on the internet inaccessible and does not permit the automatic exploitation of internet content. In SIS, data is represented with explicit semantic in logic-based syntaxes to facilitate its interpretation and processing by both humans and computers (Dieter, Frank, Michel, & Hans, 2000). The logic-based description of semantic data in SIS is carried out with languages such as RDF, RDF Schema (RDFS) and OWL (Yuang, Li, & Wang, 2013; Zhou & Yongkang, 2013).

- RDF(S) - RDF and RDFS are standard languages for representing semantic data on the internet (Dieter et al., 2000; Lu, Lei , Jean-Sébastien, Chen, Yue, & Yong, 2007). Anything on the internet is called a resource; examples of resources are web pages, emails, information retrieved from databases, web services and so forth. RDF is a data model used to describe resources on the internet, whereas, RDFS is an improved version of RDF which provides facilities for the definition of basic semantic of the data.
- OWL - OWL was developed to overcome the weak expressive power of RDF(S) (Yuang et al., 2013; Zhou & Yongkang, 2013). The expressivity of RDF(S) is enhanced by OWL with tools for describing semantic relations between data as well as constraints or restrictions on the data (Dieter et al., 2000; Lu et al., 2007).

In practice, one does not have to write semantic data in RDF(S) or OWL by hand; several software platforms exist for the automatic encoding of semantic data in RDF(S) and OWL languages. The next subsection presents the existing approaches for storing semantic data.

Approaches for Storing Semantic Data

Three approaches are used to store ontology or semantic data, namely, in-memory, native or file systems and databases (Dieter et al., 2000; HuiJun, WenGuo, & Jian, 2011). In the in-memory approach, the computer's central memory is used to store semantic data. The advantage of this approach is that it provides quick query response times with small scale semantic data. The main drawbacks of this approach are that larger semantic data are difficult to process and the stored data are not kept permanently. In fact, in this approach, the semantic data need to be loaded in the computer memory on demand; which is inefficient and time consuming.

The native storage approach uses files to store semantic data; this enables fast loading and query of semantic data (Heymans et al., 2008). Processing large scale semantic data is one of the main drawbacks of the native storage approach. Furthermore, functionalities such as query optimisation, data recovery, transaction processing, and controlled access need to be implemented separately (Heymans et al., 2008); fortunately, these drawbacks are addressed with the database storage approach. In fact, relational databases (RDB) remain the appropriate media for storying semantic data due to the maturity of the relational database technology. Therefore, the empirical part of this study focuses on storying semantic data in relational databases.

The database storage of semantic data offers many functionalities including the storage, query, reasoning and scalability (Lu et al., 2007; Yuang et al., 2013; Zhou & Yongkang, 2013). Two approaches are used to store semantic data in databases: generic and specific schema (Zhou & Yongkang, 2013). In the generic schema approach, a table is used to store semantic data in RDB (Dieter et al., 2000); the columns of the table are the elements of RDF statements of the ontology. An improved version of the generic schema approach is called normalized triple store; it uses two more tables to store semantic data with the purpose of making join queries less expensive (Hertel, Broekstra, & Stuckenschmidt, 2009).

The specific schema approach uses many tables to store semantic data; the number of tables utilized is guided by the content of the semantic data. The specific schema approach is further divided into 3 categories, namely, horizontal, vertical partitioning and hybrid approaches (Dieter et al., 2000; Zhou & Yongkang, 2013). These components of the specific schema approach use various combinations of tables to store semantic data in databases. The hybrid approach combines both vertical and horizontal approaches to store semantic data. The software platforms required to store and query semantic data are presented in the next subsection.

Software Platforms for Semantic Data Storage

To enable the storage and query of semantic data, several platforms have been developed. The most popular of these platforms are: AllegroGraph, Jena, Open Anzo, Minerva (Zhou, Ma, Liu, Zhang, Yu, & Pan, 2006) and Sesame (Fensel, Hendler, Lieberman, Wahlster, & Berners-Lee, 2005).

AllegroGaph is a server application that is accessed remotely by client applications. It enables the storage and query of semantic data and provides an API for the direct access to these data without any use of queries. Minerva is a component of the Integrated Ontology Development Toolkit; it is used as a library in Eclipse Integrated Development Environment (IDE) to store semantic data (Zhou et al., 2006). Open Anzo was developed by IBM; it can be used in three different modes to store and query semantic data: (1) embedded in an application, (2) installed as a server application and accessed remotely by clients or (3) run locally (Stegmaier, Gröbner, Döller, Kosch, & Baese, 2009). Jena API is integrated into Eclipse IDE as a library; it enables the creation and storage of semantic data in different formats (Stegmaier et al., 2009; Alamri, 2012). Sesame is a Software Development Kit (SDK) that was developed in the European IST project On-to-Knowledge (Fensel et al., 2005). It enables semantic data to be queried or exported. The abovementioned platforms for storing and querying semantic data are summarized in Table 1.

Platform	License	Operating system	Type of Storage
AllegroGraph	Commercial/Free	Linux	Native
Jena	Free/Open Source	Windows/Linux	Memory, Native, RDB
Sesame	Free/ Open Source	Windows/Linux	Memory, Native, RDB
Open Anzo	Free/ Open Source	Windows Linux	RDB
Minerva	Free	Windows/Linux	RDB

Table 1. Platforms for Storing and querying Semantic Data

The second column of Table 1 indicates whether the platform is commercial or open source. Jena API and Sesame are used in the experiments in this study as they both enable the creation/import and loading of RDF and OWL ontologies into MySQL RDBMS. Furthermore, Sesame, Jena API and MySQL are all open source platforms and are accessible free of charge with full functions and support from the Internet.

Dataset

The dataset in this study is constituted of two ontologies of the e-government domain, namely, Ontology of Development Project Monitoring (OntoDPM), and Central Government ontology (CGOV). The OntoDPM ontology is a knowledge-based model for e-government monitoring of development projects in developing countries (Fonou-Dombeu & Huisman, 2011). In fact, governments in developing countries receive aids from donors/international organizations to improve the wellbeing of their citizens through the implementation of development projects that focus on building hospitals, improving healthcare, providing education, water, electricity and so forth to the population. OntoDPM was created as a semantic data model for potential e-government applications that may be used to interface those projects for better monitoring, transparency and efficiency (Fonou-Dombeu & Huisman, 2011).

The CGOV is an ontology of the UK central government (Cgov, nd). The Friend of a Friend (FOAF) ontology (Challenger, 2012) is part of CGOV. The FOAF ontology describes the social relationships amongst people and their activities. CGOV adds the professional relationships on top of FOAF thus enabling CGOV to describe the social and professional relationships amongst government officials. In simple terms, CGOV is used to model people and the relationships between them as well as their activities in the UK central government.

Computer and Software Environment

The experiments in this study were carried out on a computer with the following characteristics: 64-bit Genuine Intel processor, Windows 8 release preview, 4 GB RAM and 160 GB hard drive. Protégé version 4.3 was installed in the computer and used to create the OWL code of OntoDPM ontology. The Apache tomcat server version 6.0 was installed to deploy the Sesame server. Jena API was configured in the Eclipse IDE version 4.2. Finally, the Wamp server was installed to enable access to MySQL backend DBMS via Sesame and Jena API.

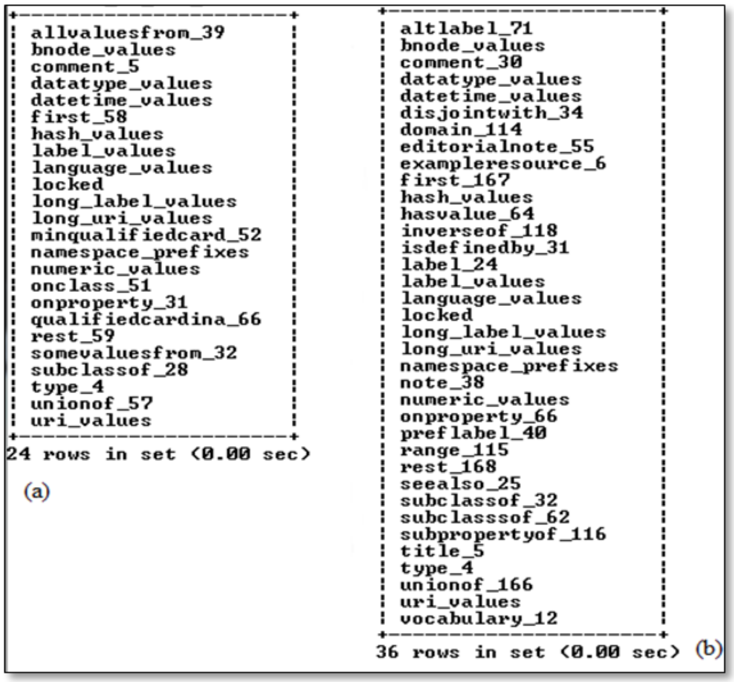
Experimental Results

This section presents the experimental results of the storage of semantic data in RDB within Sesame and Jena API platforms. The underlying database structures used by the two platforms to store semantic data are reported. In both platforms, various tables are created to stored semantic data in relational databases; it is worth noting that the tables are created based on the content of the ontology rather than the database schemas as in the relational database technology.

Storage of Semantic Data in Sesame

Let us recall that Sesame is an application that enables to store and query ontologies in relational databases. In this study, the RDBMS utilised is MySQL. For every new ontology loaded into MySQL by Sesame, 12 tables are created. Thereafter, additional tables are created based on the content of the ontology. Figure 2 shows the 24 and 36 tables that were created by Sesame to store the OntoDPM (Figure 2 (a)) and CGOV (Figure 2(b)) ontologies in MySQL databases, respectively.

The 12 default tables used by Sesame to store ontology in MySQL are listed in Table 2 along with short descriptions of their functionalities. The 12 default tables store general information on the ontology such as the data types, language, dates, access status, literals, resources, namespace and so forth. One notices that the 12 default tables listed in Table 2 are part of the tables created to store the OnDPM (Figure 2(a)) and CGOV (Figure 2(b)) ontologies in MySQL. This indicates that apart from the 12 default tables, other tables were created as well by Sesame to store the ontologies in RDB.



```
allvaluesfrom_39
bnode_values
comment_5
datatype_values
datetime_values
first_58
hash_values
label_values
language_values
locked
long_label_values
long_uri_values
minqualifiedcard_52
namespace_prefixes
numeric_values
onclass_51
onproperty_31
qualifiedcardina_66
rest_59
somevaluesfrom_32
subclassof_28
type_4
unionof_57
uri_values
24 rows in set (0.00 sec)
```

(a)

```
altlabel_71
bnode_values
comment_30
datatype_values
datetime_values
disjointwith_34
domain_114
editorialnote_55
exampleresource_6
first_167
hash_values
hasvalue_64
inverseof_118
isdefinedby_31
label_24
label_values
language_values
locked
long_label_values
long_uri_values
namespace_prefixes
note_38
numeric_values
onproperty_66
preflabel_40
range_115
rest_168
seealso_25
subclassof_32
subclasssof_62
subpropertyof_116
title_5
type_4
unionof_166
uri_values
vocabulary_12
36 rows in set (0.00 sec)
```

(b)

Figure 2. Tables Created by Sesame to Store Ontology in MySQL: (a) OntoDPM, (b) CGOV

Table Name	Description
uri_values	stores resources and literals
long_uri_values	stores resources and literals longer than 255 characters
namespace_prefixes	stores all namespaces found in the ontology
datetime_values	stores all dates and time used as values
numeric_values	stores numeric value found in the ontology.
label_values	store labels found in the ontology
long_label_values	store labels found in the ontology, longer than 255 characters
language_values	stores the languages found in the ontology
datatype_values	stores the different datatypes found in the ontology
hash_values	stores hash values generated for the uri_values data
bnode_values	stores blank nodes

Table 2. Default Tables Created by Sesame to Store Semantic Data in RDB

The extra tables were created to accommodate other constituents of the ontology. In fact, the semantic of the data is created in an ontology using concepts/classes, relationships/properties between classes and instances/literals. In Figure 2, for instance, the tables with the prefixes such as subclassof, onproperty, isdefinedby, etc. represent the relationships in the ontology. Furthermore, there are also restrictions on the elements of an ontology. A restriction is a specific condition on an element of the ontology. For instance, in Figure 2, the tables with the prefixes minqualifiedcard, qualifiedcardina, numeric, etc. are restrictions on the elements of the ontologies.

42	http://reference.data.gov.uk/def/central-government/Counsel
43	http://reference.data.gov.uk/def/central-government/MinisterialCommittee
44	http://reference.data.gov.uk/def/central-government/CabinetCommittee
45	http://www.w3.org/ns/org#OrganizationalUnit
46	http://reference.data.gov.uk/def/central-government/CivilService
47	http://www.w3.org/ns/org#Organization
48	http://reference.data.gov.uk/def/central-government/CivilServiceCommittee
49	http://reference.data.gov.uk/def/central-government/Committee
50	http://reference.data.gov.uk/def/central-government/CivilServicePost
51	http://reference.data.gov.uk/def/central-government/Post
52	http://reference.data.gov.uk/def/central-government/CorporationSole
53	http://dbpedia.org/resource/Corporation_sole
54	http://www.w3.org/ns/org#FormalOrganization
55	http://www.w3.org/2004/02/skos/core#editorialNote
56	http://en.wikipedia.org/wiki/Parliamentary_Counsel
57	http://reference.data.gov.uk/def/central-government/SeniorCivilServicePost
58	http://reference.data.gov.uk/def/central-government/Department
59	http://dbpedia.org/resource/Departments_of_the_United_Kingdom_Government
60	http://en.wikipedia.org/wiki/Departments_of_the_United_Kingdom_Government
61	http://reference.data.gov.uk/def/central-government/DeputyDirector
62	http://www.w3.org/2000/01/rdf-schema#subClasssOf
63	http://www.w3.org/2002/07/owl#Restriction
64	http://www.w3.org/2002/07/owl#hasValue

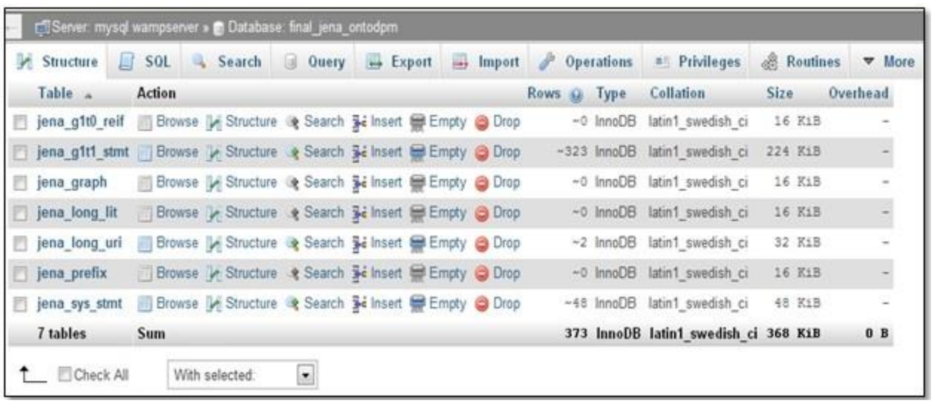
Figure 3. Partial View of the Records of the *uri_values* Table for CGOV Ontology

The classes and instances are the actual data in an ontology. They are represented as resources and literals in the ontology. In Sesame, a table called uri_values (see bottom of Figure 2 (a) and second table from the bottom of Figure 2 (b)) is used to store the literals and resources in the OntoDPM and CGOV ontologies. A partial view of the records in the uri_values table for the CGOV ontology database is depicted in Figure 3. The records of the uri_values table in Figure 3 are mainly the Uniform Resource Identifiers (URIs) of the literals and resources in the ontology. The resources and literals in the uri_values table are indexed with numbers to improve the efficiency of queries. Some examples of resources in Figure 3 include the records indexed/numbered 47, 53, 54, 56, 59 and 60. These records correspond to the concepts/classes of the CGOV ontology including organization, Corporation Sole, Formal Organization, Parliamentary Counsel and Department of the United Kingdom Government in the ontology. Also in Figure 3, some examples of literals are all URIs that contain the concepts central-government; these literals constitute the instances or branches of the UK central government.

In a nutshell, if an input ontology includes many relationships and restrictions, Sesame will create more tables in the database to store them (Figure 2). Therefore, it can be concluded that Sesame uses the specific schema approach in which the number of tables created depends on the content of the ontologies to store semantic data in RDB (Dieter et al., 2000; Lu et al., 2007; Zhou et al., 2013). The next subsection reports on the experimental results on Jena.

Storage of Semantic Data in Jena

Contrary to Sesame, Jena uses 7 tables to store an ontology in MySQL database. The number of tables do not increase after the ontology has been loaded in the RDB as in Sesame. The screenshot in Figure 4 depicts the 7 tables created by Jena to store semantic data; these tables are labelled with a Jena prefixes (left side of Figure 4).


Server: mysql wampserver » Database: final_jena_ontodpm

Structure SQL Search Query Export Import Operations Privileges Routines More

Table	Action	Rows	Type	Collation	Size	Overhead
jena_g1t0_reif	Browse Structure Search Insert Empty Drop	~0	InnoDB	latin1_swedish_ci	16 KiB	-
jena_g1t1_stmt	Browse Structure Search Insert Empty Drop	~323	InnoDB	latin1_swedish_ci	224 KiB	-
jena_graph	Browse Structure Search Insert Empty Drop	~0	InnoDB	latin1_swedish_ci	16 KiB	-
jena_long_lit	Browse Structure Search Insert Empty Drop	~0	InnoDB	latin1_swedish_ci	16 KiB	-
jena_long_uri	Browse Structure Search Insert Empty Drop	~2	InnoDB	latin1_swedish_ci	32 KiB	-
jena_prefix	Browse Structure Search Insert Empty Drop	~0	InnoDB	latin1_swedish_ci	16 KiB	-
jena_sys_stmt	Browse Structure Search Insert Empty Drop	~48	InnoDB	latin1_swedish_ci	48 KiB	-
7 tables	Sum	373	InnoDB	latin1_swedish_ci	368 KiB	0 B

Check All With selected:

Figure 4. Tables Created in Jena to Store the OntoDPM Ontology in RDB

Short descriptions of these seven tables are provided in Table 3. The tables jena_long_uri and jena_long_lit store the long URIs and literals in the ontology, respectively. The remaining (short) URIs and literals are kept in the jena_glt1_stmt table.

Table Name	Description
Jena_g1t0_reif	stores reified data
Jena_long_uri	stores all long URIs in the ontology
Jena_long_lit	stores literals longer than 255 characters
Jena_prefix	stores all prefixes in the ontologies
Jena_graph	stores data about all the ontologies loaded in the database
Jena_sys_stmt	stores system information on the database
Jena_g1t1_stmt	Stores all statements in the ontology

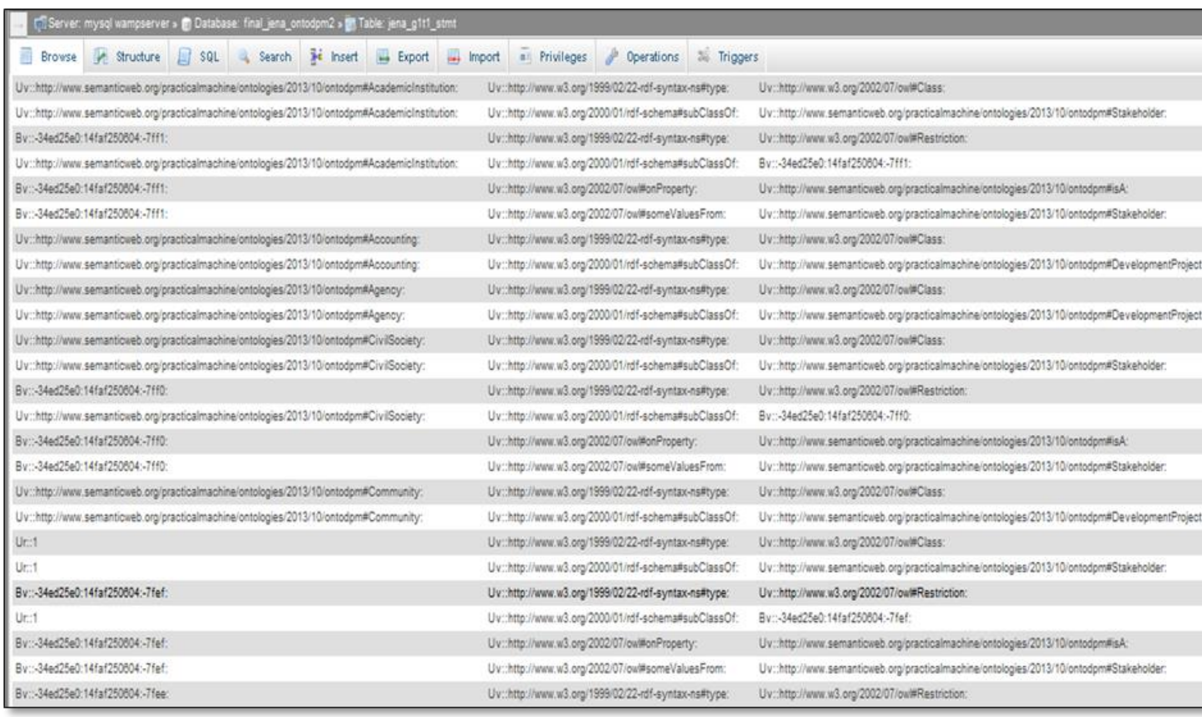
Table 3: Tables Created in Jena to Store Semantic Data in RDB

The screenshot in Figure 5 depicts a partial view of the jena_glt1_stmt table for the OntoDPM ontology stored in MySQL. The records of the tables are short URIs of the OntoDPM ontology content including resources, literals and restrictions. Furthermore, each record of the jena_glt1_stmt table has 3 columns; these columns correspond to the elements of the RDF statements in the OntoDPM ontology.

After analysing the tables of the databases of the OntoDPM and CGOV ontologies created with Jena in MySQL, it appeared that these databases have the same number of tables (Figure 4 and Table 3). Therefore, it can be concluded that Jena uses the specific schema approach in which all the databases used to store semantic data have the same number of tables. This is contrary to Sesame where the number of tables of ontology databases depend on the content of the ontologies (Zhou et al., 2013).

CONCLUSION

This study has addressed the topic of semantic data storage in IS. It involved a review of the existing approaches, languages and software platforms for storing ontology or semantic data. Thereafter, the design research was applied to conduct experiments with two ontologies of the e-government domain on two popular platforms for storing and querying semantic data, namely, Sesame and Jena API. The analysis of the experimental results revealed the underlying structures used by the two platforms to store semantic data in relational databases. Precisely, the results showed that Sesame uses the specific schema approach in which the number of database tables created to store the semantic data depends on its content, whereas, Jena utilises the specific schema approach in which all the databases of ontologies have the same number of tables. The empirical findings of the study shed light on the technologies and approaches utilised to store semantic data in relational databases. This may contribute to the understanding of semantic technologies in IS and foster the development of semantic information systems.



Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#AcademicInstitution:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#AcademicInstitution:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Bv::-34ed25e0:14faf250604:-7ff1:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Restriction:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#AcademicInstitution:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Bv::-34ed25e0:14faf250604:-7ff1:
Bv::-34ed25e0:14faf250604:-7ff1:	Uv::http://www.w3.org/2002/07/owl#onProperty:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#isA:
Bv::-34ed25e0:14faf250604:-7ff1:	Uv::http://www.w3.org/2002/07/owl#someValuesFrom:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Accounting:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Accounting:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#DevelopmentProject:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Agency:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Agency:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#DevelopmentProject:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#CivilSociety:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#CivilSociety:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Bv::-34ed25e0:14faf250604:-7ff0:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Restriction:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#CivilSociety:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Bv::-34ed25e0:14faf250604:-7ff0:
Bv::-34ed25e0:14faf250604:-7ff0:	Uv::http://www.w3.org/2002/07/owl#onProperty:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#isA:
Bv::-34ed25e0:14faf250604:-7ff0:	Uv::http://www.w3.org/2002/07/owl#someValuesFrom:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Community:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Community:	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#DevelopmentProject:
Ur::1	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Class:
Ur::1	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Bv::-34ed25e0:14faf250604:-7fef:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Restriction:
Ur::1	Uv::http://www.w3.org/2000/01/rdf-schema#subClassOf:	Bv::-34ed25e0:14faf250604:-7fef:
Bv::-34ed25e0:14faf250604:-7fef:	Uv::http://www.w3.org/2002/07/owl#onProperty:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#isA:
Bv::-34ed25e0:14faf250604:-7fef:	Uv::http://www.w3.org/2002/07/owl#someValuesFrom:	Uv::http://www.semanticweb.org/practicalmachine/ontologies/2013/10/ontodpm#Stakeholder:
Bv::-34ed25e0:14faf250604:-7fee:	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type:	Uv::http://www.w3.org/2002/07/owl#Restriction:

Figure 5. Partial View of *jena_g1t1_stmt* Table for the OntoDPM Ontology

REFERENCES

Alamri, A. (2012) The relational database layout to store ontology knowledge base, *Proceedings of the International Conference on Information Retrieval & Knowledge Management (CAMP)*, Kuala Lumpur, Malaysia, 13-15 March, pp.74-81.

Cgov (nd): Central Government Ontology. An Ontology of the UK Central Government; Retrieved August 24, 2017 from http://reference.data.gov.uk/def/central-government

Challenger, M. (2012) The Ontology and Architecture for an Academic Social Network, *International Journal of Computer Science Issues*, 9, 1, 22-27.

Chebotko, A., Deng, Y., Lu, S., Fotouhi, F. and Aristar, A. (2005) An Ontology-Based Multimedia Annotator for the Semantic Web of Language Engineering, *International Journal on Semantic Web and Information Systems (IJSWIS)*, 1, 1, 50-67.

Dieter, F., Frank, V.H., Michel, K. and Hans, A. (2000) On-To-Knowledge: Ontology-based Tools for Knowledge Management, *In Proceedings of the eBusiness and eWork (EMMSEC 2000) Conference*, Madrid, Spain,18-20, pp. 1-7.

Fensel, D., Hendler, J., Lieberman, H., Wahlster, W. and Berners-Lee, T. (2005) Sesame: An Architecture for Storing and Querying RDF Data and Schema Information, Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential, *MIT Press*, 197-222.

Fernandez, J.D., Arias, M., Martinez-Prieto, M., Gutierrez, C. (2013) Management of Big Semantic Data, Big Data Computing, Taylor and Francis/CRC Press, Rajendra Akerkar (eds.), 131-167.

Fill, H.G. (2009a) Visualization of Semantic Information Systems, *Gabriel Gurley (eds.)*, Germany.

Fill, H.G. (2009b) Design of Semantic Information Systems using a Model-based Approach, *Proceeding of AAAI Spring Symposium*, CA, USA, March 23–25, pp. 19-24.

Fonou-Dombeu, J.V. and Huisman, M. (2011) Combining Ontology Development Methodologies and Semantic Web Platforms for E-government Domain Ontology Development, *International Journal of Web & Semantic Technology (IJWesT)*, 2, 2, 12-25.

Guido, A. L. and Paiano, R. (2010) Semantic Integration of Informations Systems, *International Journal of Computer Networks & Communication (JCNC)*, 2, 1, 48-64.

Hevner, A. R. and Chatterjee, S. (2015) Design Research in Information Systems,

Hertel, A., Broekstra, J. and Stuckenschmidt, H. (2009) RDF Storage and Retrieval Systems, International Handbooks on Information Systems, 489-508.

HuiJun, D., WenGuo, W. and Jian, Y. (2011) B-Tree Algorithm for Partitioned Index based on CIDR List, *Proceedings of the International Conference of Information Technology, Computer Engineering and Management Sciences*, Nanjing, Jiangsu, China, 24-25 September, pp. 124-128.

Heymans, S., Ma, L., Anicic, D., Ma, Z., Steinmetz, N., Pan, Y., Mei, J., Fokoue, A., Kalyanpur, A. Kershenbaum, A., Schonberg, E., Srinivas, K., Feier, C., Hench, G. Westzstein, B. and Keller, U. (2008) Ontology Reasoning with Large Data Repositories, *Ontology Management*, Martin Hepp, Pieter De Leenheer, Aldo De Moor, York Sure (eds.), Springer, 7, 89-128.

Lili, X., Lee, S. and S. Kim, S. (2010) E-R model based RDF data storage in RDB, *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, Chengdu, China, 9-11 July, pp. 258-262.

Lu, J., , M., Lei , Z., Jean-Sébastien, B., Chen, W., Yue, P. and Yong, Y. (2007) SOR:A practical System for Ontology Storage, Reasoning and Search, *Proceedings of the 33rd international conference on very large databases*, Vienna, Austria, September 23-28, pp.1402-1405.

Nalepa, G.J. and Furmanska, W.T. (2009) Review of semantic web technologies for GIS, Automatyka: półrocznik Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie, 13, 2, 485– 492.

Pinto, H. S., Staab, S. and Tempich, C. (2004) DILIGENT: Towards a Fine-Grained Methodology for Distributed, Loosely-controlled and Evolving Engineering for Ontologies, *Proceedings of the 16th European Conference on Artificial Intelligence, Valencia, Spain*, 22-27 August, pp. 393-397.

Prt, (nd) Parrot: a RIF and OWL documentation service; Retrieved August 24, 2017 from http://ontorule-project.eu/parrot.html

Ramanujam, S., Gupta, A., Khan, L., Seida, S. and Thuraisingham, B. (2009). R2D: A Bridge between the Semantic Web and Relational Visualization Tools, *Proceedings of the 3rd IEEE International Conference on Semantic Computing (ICSC)*, Berkeley, CA, USA, 14-16 September, pp. 303-311.

Rajbhandari, P., Gosai, R., Shah, R.C. and Pramod, K.C. (2012) Semantic Web in Medical Information Systems, *International Journal of Advanced in Engineering & Technology*, 5, 1, 536-543.

Ricca, F., Grasso, G., Liritano, S., Dimasi, A., Lelpa, S.M., Manna, M. and Leone, N. (2010) A Logic-Based System for e-Tourism, *Fundamenta Informaticae*, 105, 35–55.

Soualah-Alila, F., Faucher, C., Bertrand, F., Coustaty, M. and Doucet, A. (2015) Applying Semantic Web Technologies for Improving the Visibility of Tourism Data, *Proceedings of the Eighth Workshop on Exploiting Semantic Annotations in Information Retrieval*, October 23, Melbourne, VIC, Australia, pp. 5-10.

Stegmaier, F., Gröbner, U., Döller, M., Kosch, H. and Baese, G. (2009) Evaluation of Current RDF Database Solutions, *Proceedings of the 10th International Workshop on Semantic Multimedia Database Technologies in conjunction with the 4th International Conference on Semantic and Digital Media Technologies (SAMT 2009)*,Graz, Austria, 2-4 December, pp.1-17, 2009.

Sure, Y., Staab, S. & Studer, R. (2002) Methodology for development and employment of ontology based knowledge management applications, *ACM SIGMOD Record Journal*, 31, 4, 8-23.

Suarez-Figueroa, M.C. (2010) NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse, *PhD thesis*, Universidad Politecnica de Madrid, Spain.

Yadagiri, N. and Ramesh, P. (2013) Semantic Web and the Libraries: An Overview, *International Journal of Library Science*, 7, 1, 80-94.

Yuang, Y., Li, G. and Wang, X. (2013) Semantic Web Rough Ontology: Definition, Model and Storage Method, *Proceedings of the 6th International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII)*, Xi'an, China, 23-24 November, pp.104-107.

Vaishnavi, V. and Kuechler, W. (2004) Design Research in Information Systems; Retrieved September 28, 2017 from http://desrist.org/desrist/content/design-science-research-in-information-systems.pdf

Wilkes, J., Hoover, C., Keer, B., Mehra, P. and Veitch, A (2006) Storage, data, and information systems, *Hewlett-Packard Laboratories*, USA.

Zhou, J., Ma, L., Liu, Q., Zhang, L., Yu, Y. and Pan, Y. (2006) Minerva a scalable owl ontology storage and inference system, *Proceedings of the first Asian semantic web conference*, Beijing, China, 3-7 September, pp 429-443.

Zhou, Z. and Yongkang, X. (2013) A study on ontology storage based on relational database, china, *IEEEConference Anthology*, 1-8 January, pp.1-5.