Building a New Semantic Social Network Using Semantic Web-Based Techniques

Ahmed M. Daoud [1],*, Khlid M. Hosny [2] and Ehab R. Mohamed [3]

[1] Department of Information Technology, Faculty of Computers and Information, Zagazig University, Zagazig 44519, Egypt; developerdaoud@gmail.com

[2] Department of Information Technology, Faculty of Computers and Information, Zagazig University, Zagazig 44519, Egypt; k_hosny@yahoo.com

[3] Department of Information Technology, Faculty of Computers and Information, Zagazig University, Zagazig 44519, Egypt; ehab.rushdy@gmail.com

* Correspondence: developerdaoud@gmail.com

Abstract

Most people are more or less related to the web by participating in a kind of social networking site. Nowadays, Semantic Web technology is playing a crucial role in these sites as they contain an enormous amount of data about persons, pages, events, places, corporations, etc. This research is a Semantic Web application designed to create a new semantic social community called Socialpedia. It links the already existing social public information to the newly public ones. This information is linked with different pieces of information on the web to construct a new immense data container. The resulting data container can be processed using a variety of Semantic Web techniques to produce machine-understandable content. This content shows the promise of using integrated data to improve Web search and Web-scale data analysis, unlike conventional search engines or social ones. This community involves obtaining data from traditional users known as contributors or participants, linking data from existing social networks, extracting structured data in triples using predefined ontologies, and finally querying and inferring such data to obtain meaningful pieces of information. Socailpedia supports all popular functionalities of social networking websites besides the enhanced features of the Semantic Web, providing advanced semantic search that acts as a semantic search engine.

Keywords: Ontology, OWL, RDF, R2RML, SPARQL, Semantic Web (SW), Semantic Search, Social Networking Sites (SNSs)

1. Introduction

The Semantic Web, called Web 3.0 or Web of Data, "SW is not technology, it is philosophy" [1]. It has been designed to allow for the integration of data across a variety of applications, Web sites and services, and software systems to ensure interoperability. It applies advanced knowledge technologies to fill the knowledge gap between humans and machines [2]. It also enables both users and machines to share and combine information easily and therefore allows machines to intelligently access different data sources. Thus, it can greatly boost the computer's ability to provide more relevant answers. This means providing knowledge in forms that computers can readily process and reason with [3] and therefore participate in accomplishing manual tasks automatically. The Resource Description Framework "RDF" [4] is a popular, easily accessible, and queryable format for representing data modeled as a graph. In that role, RDF is considered the basic pillar of constructing the Semantic Web. It is a graph data model standardized by W3C for encoding and processing metadata and it consists of statements called triples. A collection of triples constructs an RDF graph in which each triple represents a piece of knowledge for describing a resource and consists of a subject (resource), predicate (property), and object (value). These triples can be used for further processing and constructing a web of data. It is defined as a collection of information linked in a way so that it can be easily processed by machines. It is queried by The SPARQL Protocol and RDF Query Language "SPARQL" [5], the standard query language specially designed to query RDF data stores. A significant advantage of this language is that we do not need to know the way, how the data are stored in the queried source. We only need to know what kind of entities are present in the source and what properties we can find on them. However, that can be also retrieved by a query so we can query the source with no more knowledge than how to connect to the endpoint. The Web Ontology Language "OWL" [6] is a recommended language, by W3C, for modeling and expressing ontologies. An ontology defines shared terms/vocabulary in some domain knowledge and can be expressed in formal languages that are specifically designed for a web-based scenario and have well-defined logic-based semantics. It provides machines with a greater ability to interpret Web content using its rich vocabulary and underlying formal semantics of Description Logics. Description Logics (DL) [7] are logics that are designed to represent and reason on structured knowledge. They are considered fragments of First-Order logic (FOL). Therefore, ontologies are logical theories.

A variety of research studies have focused on strategies for crawling data from Social Networking Sites (SNSs). Crawler agents for Facebook, Twitter, etc. are used to crawl user profiles. Crawling data is not the best choice for collecting data from these large-scale corporations. This may have several reasons such as complexity, scalability, the computational resources needed, the time required for mining data, and the crawling overhead that was up to 44 terabytes in 2010 [8], so many research are based on sampling from a starting seed profile for data analysis [9]. Moreover, Social network companies like Facebook are reluctant to share their data for research purposes. So, they continuously upgrade their terms, conditions, and privacy policies. This way these crawling techniques, over time, become unusable. For all these reasons, the best choice is to integrate partially publicly available data [10] from these corporations. For this purpose, they only provide graph API for querying user's public data.

Accessing and utilizing enterprise or Web data that is scattered across multiple extremely heterogeneous data sources is an important task for both applications and users. Many of these data sources are very dynamic since a large amount of the newly collected data is continuously available on the Web [11, 12]. The main emphasis here, however, is on the integration of the data and subsequent processes rather than on how the data are collected. This motivates the authors to construct a new open privacy community, called Socialpedia, which acts as a central interlinking hub to all user data on the web. The idea is based on integrating multiple existing social media data with new user data to construct one large social data community. Exploiting the Semantic

Web technologies, this community will be helpful for information searching, job seeking, consistency checking, inferencing, and running more complicated queries unlike traditional search engines or social ones.

The remaining of this paper is structured into five sections: Section 2 presents a summary of the most related work. Section 3 describes the methodology used to conduct this work. Section 4 discusses the experiments of the proposed method. The most important results are summarized in Section 55. Lastly, Section 6 summarizes this work and addresses some future extensions of it.

2. Related Work

The transformation process of the relational database "RDB" schema and data into OWL and RDF respectively are called RDB2RDF. Over the past few years, several RDB2RDF tools have been developed for this mapping process [13, 14]. Also, some methods for ontology scheme representation have been created [15-22]. Currently, there are two possible approaches (and both of them are published as a W3C Recommendation).

The Direct Mapping approach [23] is a simple transformation of the relational database to the RDF (so it creates the RDF resources according to the table names, their columns, and other information like primary keys, foreign keys, etc.). The approaches [15, 21] proposed an automatic mapping where the relational database schema is directly mapped into ontology "represented in RDF(S)-OWL" and the data into instances "represented in RDF triples" based on the structure of the database schema. While in approach [20], allows easy creation of a database to ontology involving basic one-to-one correspondence between ontology and database schema mappings in a compact notation. It also translates the resulted mappings into executable D2RQ mapping to create a SPARQL endpoint over the conceptual data from the database. The survey [24] showed a set of approaches to convert the implicit RDB schema into explicit semantics in form of ontology and the data into RDF triples.

The other approach is more complex than direct mapping, called R2RML [25] mapping. It is using a mapping definition (typically in the form of a turtle [26] RDF syntax file) that is manually created and that defines the exact way how to map the relational data into the RDF form. Some early tools have been developed such as Triplify [27], D2R Server [28], ODEMapster [29], and OpenLink Virtuoso [30]. These tools enable experts to create the mapping between RDB content and RDF based on a target ontology. The survey [24] revealed that these tools adopt a proprietary mapping language. Also, Some R2RML-based tools were developed such as DB2Triples [31] for both DM and R2RML implemented in java, Morph-RDB [32] is a java and scala implementation for R2RML, Oracle Database 19c [33] is a commercial tool that supports graph analytics to enable data storage, access, and analysis in an Oracle database and supports both DM and R2RML mapping, RDF-RDB2RDF [34] is an open-source library implemented in Perl, Virtuoso RDF Views [30] is an open-source and commercial Web interface that supports R2RML only, XSPARQL [35] is a command-line tool implemented in java with R2RML supports, Ultrawrap [36, 37]. It is a commercial tool that supports DM, R2RML, and D2RQ mapping. The implementation report of these DM and R2RML tools is shown in [38] and their compliance test cases are shown in [39]. Recently, many platforms developed to deal with Semantic Web-linked data to facilitate the query process for non-expert users. The authors in [40] developed software to facilitate and automate the process of creating the mapping file. Their tool enables experts and semi-experts to create and edit the mappings using GUI wizards. This reduces the errors and validates the resulting mapping.

SQL and SPARQL are similar since they offer to users the possibility of accessing, creating, combining, and consuming structured data. The direct comparison between them is rather difficult [41]. The relational tables in databases are queried using SQL, while the semantic data triples (subject, predicate, and object) in RDF stores are queried using SPARQL. Although Triple Stores are a better fit than RDBMS, the latter is wider spread and usually contains more data. Triple Stores and SPARQL are built to handle undetermined and varying or growing data representations. In contrast, RDBS and SQL require a well-designed and complete database layout. Varying or additional relations that require subsequent changes to the database layout are costly. Furthermore, SPARQL naturally interacts better with ontologies than SQL, since the graph-based topology of ontologies is similar to the data representation of many Triple Stores. The mapping of SQL queries to SPARQL queries and vice versa becomes a relevant need. Several research works have been recently developed regarding this transformation. The authors in [42] proposed an approach to interrogate RDF data using SQL queries. They developed an algorithm that translates both simple and complex queries (without join) into their equivalent SPARQL ones. Soussi and Bahaj [43] suggested an efficient algorithm to convert SQL queries containing Left/Right Outer Join command(s) (simple and nested form) to their semantically equivalent SPARQL ones with Optional pattern(s) (simple and nested). Their proposed algorithm is beneficial for those working with relational data and aiming to interact with semantic databases without migration of their system and training their users in the semantic technologies new for them. Besides, this algorithm is beneficial in terms of complexity and query execution time as it replaces the heavy and costly SQL joins with SPARQL Optional clauses. The Ontop [44, 45] system was designed to query data stored in RDBs, with the vocabularies of the data and OWL2QL ontologies. The authors of [46] developed a SPARQL query processing engine called Quest, the core of the Ontop, to convert the SPARQL queries against RDF graph and ontology into SQL queries against RDB. The ontop system provides a fast conversion from SPARQL to SQL [47] typically (4–15ms) and efficient optimization of SQL queries. It can run 44k queries per hour over 200 million triples. The authors of [48] implemented a procedure to answer SPARQL queries under the OWL2QL semantics where the data are stored in RDB with schema connected to OWL2QL ontology using R2RML. The authors of [49] proposed an efficient approach to generate SQL queries from SPARQL queries based on combining the logic programming optimization techniques with the SQL optimization fields. This approach provides a specification of the SPARQL semantics used for the translation process and supports R2RML mapping over RDB schemas.

Concerning social networks, many studies were conducted to analyze, extract, link, and convert their data into useful data models. The Friend of a Friend “FOAF” [50] is a widely accepted Semantic Web standard used to represent social networks. Many large Social Networking Sites “SNSs” use it to produce Semantic Web profiles for their users. It is a predefined dictionary that has a set of terms/vocabulary used as an ontology to describe people and the things they do. It is written in RDF/OWL and designed to be reusable alongside other ontologies to describe people, groups, companies, etc. It defines classes and their properties as well as the relationships between them. It utilizes several features of OWL so interesting inferences can be done. Using FOAF semantics and other Semantic Web techniques, a significant percentage of user profiles across multiple online social networks can be integrated [51]. And using reasoning to merge a large set of data instances, distributed among many independent websites, and to find the relationships between these data. Their study is only applied to social networks that represent their user profiles in FOAF.

There are many techniques designed to extract social information on the web. These techniques include parsers, scrapers, and APIs to access network data. Flink [52] is a system that uses some of these approaches and others to extract, analyze (including merging), and visualize social networks on the web. The scope of this work is to look only at APIs to aggregate user's public profiles from various social networks. Jie et al. [53] developed the Arnetminer system, a semantic-based social network, to extract researcher's profiles from the Web and disambiguate those having the same name. It extracts the researcher's basic, contact, and educational information and integrates his publication information from DBLP which helps in expert searching. This method used the tagging basis to collect the researcher's data from the Web. This system used a classifier based on the tagging basis to collect the researcher's data from the relevant Web pages.

3. Methodology

A lot of new technologies are used to handle data storage using non-relational databases. Although large corporations like (Facebook, Twitter, Google, etc.) started using them, most companies are still storing their data in RDBMSs. The majority of web information is currently stored in relational databases since they were dominant for the past several decades due to their simplicity and performance in managing data. We start building the user interface of our application, Socialpedia, to allow users to easily interact with it. The application is designed to allow the user to partially register some of his information. Later, one can add more information as desired. The proposed application is built upon four main phases:

1) *Integration and Storage Phase:* In this phase, allow users to link their social profiles with various social media data sources like Facebook, Twitter, Linked-in, etc. They optionally choose which social media data sources to link with. When they choose to link to one source, they are redirected to the authentication page. This tells the source that the Socialpedia application wants to get access to some information. The user needs to grant access to get his access token. Using the access token and the source API, Socialpedia could get the user's public information from these sources and store them in the database. The most relevant attributes in these sources include, but are not limited to, id, email, telephone, first name, last name, date of birth, city, location, country, work, education, relationship, religion, about, gender, etc. This process is reversible, meaning that the user can unlink his profile and all data relating to that profile are logically deleted. The architecture of this phase can be seen in Figure 1.

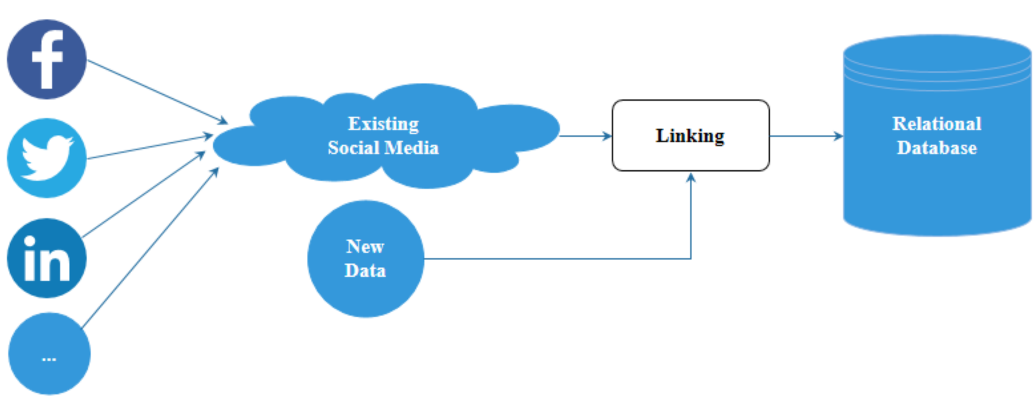

Figure 1. Storage and integration phase.

2) *Custom Mapping Phase:* the main purpose of this phase, shown in Figure 2, is to generate a mapping document that is later provided as an input for a dumping engine or any other R2RML processor to produce an RDF triples graph. This process uses the relational database tables that contain the previously linked attributes and a set of predefined ontologies such as FOAF, vCard, etc. The output is an R2RML document that contains the mapping graph, which is a representation of the R2RML mapping written in the Turtle RDF recommended syntax (.ttl), and it is used to describe the relationship between the relational database and ontology.

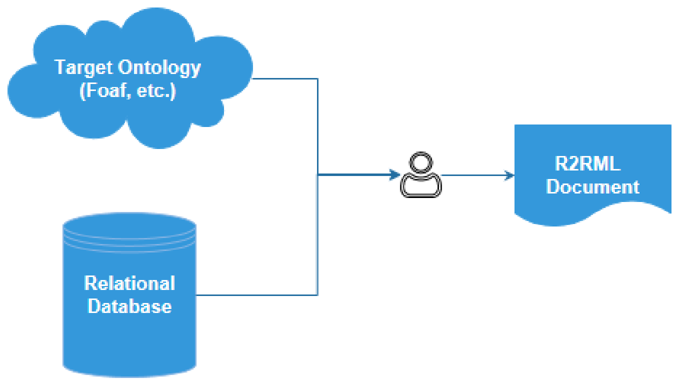

Figure 2. Custom mapping phase

3) *Dumping Phase:* this phase includes the conversion of data into semantically structured data (triples) using the R2RML mapping document from phase 2, the system database, and a selected output format (.nq in this case). We implemented a dumper engine, shown in Figure 3, to take this information as an input and produce the RDF graph document. The dumping process can be accomplished using one of the following three approaches:

I. *Persistent Automated Approach (change-based flow):* on every small change in the database, create a large dump of RDF data and load it into memory to create a SPARQL endpoint over an RDF dataset based on a predefined ontology, R2RML document, in this case, thus allows the user to run SPARQL queries against the up-to-date dump at any time. The main drawback of this approach is resource-wasting results from server workflow/bottleneck as the database is continuously changing by the users and the dump data storage size that is continuously increased each time a new data is inserted.

II. *Persistent Scheduled Approach (schedule-based flow):* this approach is based on scheduling the dump process with a specified task that runs according to a scheduled time, e.g. every certain period of day (s) at a specified time. This can utilize the server resources instead of the continuous dump process. The main drawback here is that the dump is not always up-to-data for the user to run SPARQL queries and the dump data storage size is still a problem.

III. *Just-in-Time or Instant Approach (query-time-based flow):* this approach is based on creating a virtual SPARQL endpoint without dumping any RDF data. This endpoint has a predefined ontology, the R2RML document is this case, which specifies how to represent the relational data. Every SPARQL query is converted into a SQL query, executed against the database and the result is converted back to the expected SPARQL query result. This approach is a little time-consuming due to the conversion process of queries, but there is no need to dump all data from the RDBs or even storing any RDF data and also could utilize the efficient query optimization services of the SQL engine.

Each approach has its features and drawbacks, choosing the right approach is beyond the scope of this work. The second approach is used in this paper as the user's public data are not instantly changed.

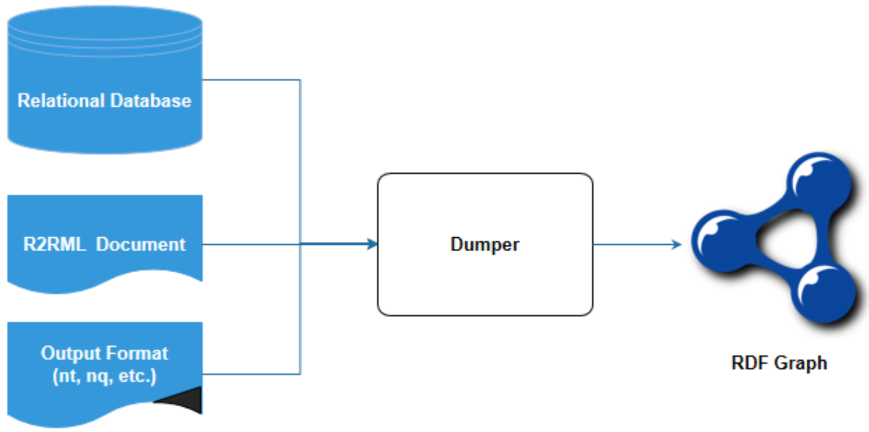

Figure 3. Dumping phase

4) *Retrieval/query phase:* as the application data increase, the resulting RDF graph is going very large. So, we also developed a splitter engine to split the resulting RDF graph from phase 3 into smaller RDF graphs suitable to be updated to the Virtuoso server. This server has a SPARQL endpoint and a faceted RDF browser that the user can use to query the RDF data. The architecture of this phase can be seen in Figure 4.

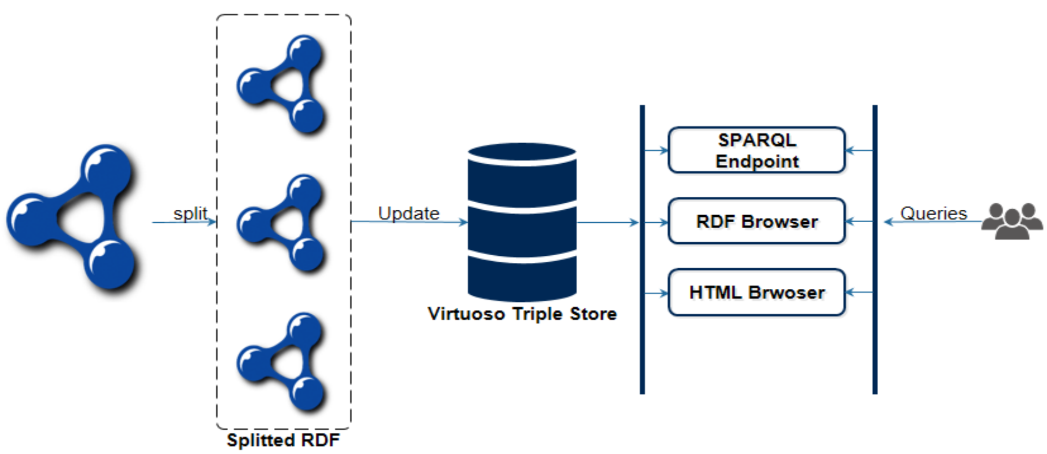

Figure 4. Retrieval and query phase

4. Experiments

The C# 4.5 (currently the newest version of the major language for the .NET Framework) is selected as the programming language of the proposed system. It is a modern language that enables using of various paradigms - imperative, declarative, generic, object-oriented, and event-driven programming. Third-party libraries are used to work with the RDF data. That means the dotNetRDF library for working with RDF and SPARQL. And the r2rml4net library for loading the mapping file. These two are the major libraries for their purposes on the .NET platform. As a relational database, MSSQL 2016 is chosen as a standard database used in .NET applications. To create a web application that will host the SPARQL endpoint we have chosen the ASP.NET MVC 5. That is a standard approach when building .NET web applications. The performance is tested on a Windows 8 64-bit laptop having 8GB of RAM and a 1.8GHz Intel Core i7-4500u CPU.

The dump process could be accomplished either manually by the author or automatically by scheduling it according to a pre-specified time. To automate the process, a windows service could be used to dumb the database according to a predefined R2RML document.

For the test purpose, the splitter engine is tested against an instance file containing 3,854,286 triples and with a 512 MB file size. It took a few seconds (typically 3 - 5s) and resulted in 65 files, 64 of them have 60000 triples and the remaining one has 14286 triples. These files are updated to the virtuoso server. The result is considered a graph G = (V, E) where V represents a finite set of vertices and E represents a finite set of edges such that $E \subseteq V \times V$.

subject	predicate	object
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.w3.org/2006/vcard/ns#given-name	"ahmed"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.w3.org/2006/vcard/ns#family-name	"daoud"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.w3.org/2006/vcard/ns#title	"Teaching assitant"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.w3.org/2006/vcard/ns#country-name	"Egypt"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.w3.org/2006/vcard/ns#homeAdr	"Kotur, Al Gharbiyah, Egypt"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/birthday	"1992-06-04"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/gender	"Male"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/homePage	"http://www.adaoud.com/"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/img	http://www.spedia.com/Images/Profiles/ahmeddaoudzsufu
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/knows	http://www.spedia.com/people/contact/1571b33d-54c9-4b
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/knows	http://www.spedia.com/people/contact/5aeb9a8c-71ed-40
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/knows	http://www.spedia.com/people/contact/85814e90-cad4-44
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/knows	http://www.spedia.com/people/contact/a1f27b66-823e-41
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/phone	"+201015950189"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/status	"married"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://xmlns.com/foaf/0.1/workplaceHomepage	"http://www.fci.com"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.spedia.com/people/contact/profileLink	"http://localhost:61167/ahmeddaoudzsufuXIW"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.spedia.com/people/contact/bio	"TA at it department"
http://www.spedia.com/people/contact/9625e72c-bdc2-4a8e-9493-ec316a66152e	http://www.spedia.com/people/contact/education	"B.S. in Infromation Technology from FCI at ZU Unive

Figure 5. Sample of RDF triples

Figure 5 displays a sample of the RDF triples graph generated using the custom mapping document, split using the splitter, and updated to the Virtuoso server. The triples are formed using subject, predicate, and object maps. By default, all RDF triples are placed in the default dataset graph.

5. Results and Discussion

Using the faceted browser's search and find service found in the OpenLink Virtuoso server, one can make text search, entity label lookup, or entity URI lookup. Some users would prefer to use the SPARQL endpoint to write SPARQL queries on their own. It is considered a query language with the capability to search graph patterns shown in Figure 6. These queries are executed against RDF graphs. It extracts information in the form of URIs, plain text, blank nodes, and typed literals. It can also extract RDF subgraphs and build new RDF graphs using information from the queried graphs. The (subject, p, o) is a pattern for what we are looking for with p and o as “unknowns”.

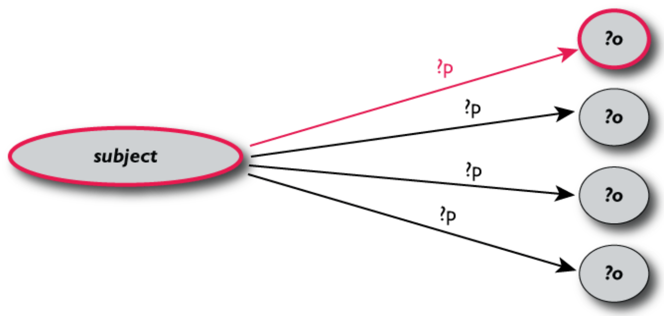

Figure 6. Graph pattern representation

Table 1 shows some sample queries that can be run against the RDF graph. The query results can be returned in many formats such as XML, JSON, RDF, or HTML. Each query results in a set of returned values, so only a snippet is included in the table.

Table 1. SPARQL queries, their meaning, and result snippets

<table>
<tr><th>Query</th><th>Meaning</th><th>Result snippet</th></tr>
<tr><td>SELECT ?name ?email
WHERE {
?person foaf:name ?name .
?person foaf:mbox ?email .
?person v:workAdr "Zagazig".
}</td><td>Find the names and emails of all the people that have names, email addresses, and work at Zagazig University.</td><td><table><tr><th>name</th><th>email</th></tr><tr><td>"ahmed daoud"</td><td>"developerdaoud@gmail.com"</td></tr><tr><td>"khalid hosny"</td><td>"k_hosny@yahoo.com"</td></tr></table></td></tr>
<tr><td>SELECT ?name ?title ?bio
WHERE {
?person foaf:name ?name.
?person v:workAdr "Zagazig".
?person sp:interests ?i.
?person v:title ?title.
?person sp:bio ?bio.
FILTER(bif:contains(?i, "programming")) .
}</td><td>Find the name, title, and bio of all people who work in Zagazig university and interested in programming.</td><td><table><tr><th>name</th><th>title</th><th>bio</th></tr><tr><td>"ahmed daoud"</td><td>"Teaching assitant"</td><td>"TA at it department"</td></tr><tr><td>"khalid hosny"</td><td>"Professor"</td><td>"professor at it department"</td></tr></table></td></tr>
</table>

The proposed approach results in merging profiles of the same person from multiple social networking websites and creating a large, unified social network from sub-networks that evolved independently. This approach could be linked, using extraction methods or APIs, to the online databases from other academic social networks such as IEEE Explore, DBLP, ACM DL, etc. [54].

Also, it could utilize the approaches used for natural language interpretation to convert natural language queries into efficient SPARQL queries [55]. The conversion process is based on ontology and uses SKIN rules and SKOS lexicon.

6. Conclusions

In this paper, a new semantic social network application, called Socialpedia, is proposed. It acts as a central interlinking hub to all user data on the web. It integrates multiple existing social media data with new user data to construct one large social data community. This application is helpful for information searching, job seeking, consistency checking, inferencing, and running more complicated queries unlike traditional search engines or social ones. This community scale could be extended to include linking to scientific web resources or any other resources to construct a titanic data container acting as an interlinking hub for all contributor's data on the web. Once done, the application provides a REST API to help other applications consume the resulted records using an OAuth 2.0 authentication protocol and supporting data exchange in XML/JSON format.

References

[1] Berners-Lee T., Hendler J., and Lassila O., "The Semantic Web," Scientific American, vol. 284, pp. 34-43, 2001.

[2] Mika P., "Social networks and the semantic web," in Proceedings of the 2004 ieee/wic/acm international conference on web intelligence, 2004, pp. 285-291.

[3] Shadbolt N., Berners-Lee T., and Hall W., "The semantic web revisited," IEEE intelligent systems, vol. 21, pp. 96-101, 2006.

[4] Manola F., Miller E., and McBride B., "RDF 1.1 primer, W3C working group note," 2014.

[5] Prud'hommeaux E. and Seaborne A., "SPARQL query language for RDF, W3C Candidate Recommendation (2006)," http://www.w3.org/TR/rdf-sparql-query/, 2007.

[6] Motik B., Patel-Schneider P. F., Parsia B., Bock C., Fokoue A., Haase P., et al., "OWL 2 web ontology language: Structural specification and functional-style syntax," W3C recommendation, vol. 27, p. 159, 2009.

[7] Baader F., Calvanese D., McGuinness D., Patel-Schneider P., and Nardi D., The description logic handbook: Theory, implementation and applications: Cambridge university press, 2003.

[8] Gjoka M., Kurant M., Butts C. T., and Markopoulou A., "Walking in facebook: A case study of unbiased sampling of osns," in 2010 Proceedings IEEE Infocom, 2010, pp. 1-9.

[9] Catanese S. A., De Meo P., Ferrara E., Fiumara G., and Provetti A., "Crawling facebook for social network analysis purposes," in Proceedings of the international conference on web intelligence, mining and semantics, 2011, pp. 1-8.

[10] Ferrara E., De Meo P., Fiumara G., and Baumgartner R., "Web data extraction, applications and techniques: A survey," Knowledge-based systems, vol. 70, pp. 301-323, 2014.

[11] Dong X. L. and Srivastava D., "Big data integration," in 2013 IEEE 29th international conference on data engineering (ICDE), 2013, pp. 1245-1248.

[12] Duggan J., Elmore A. J., Stonebraker M., Balazinska M., Howe B., Kepner J., et al., "The bigdawg polystore system," ACM Sigmod Record, vol. 44, pp. 11-16, 2015.

[13] Sahoo S. S., Halb W., Hellmann S., Idehen K., Thibodeau Jr T., Auer S., et al., "A survey of current approaches for mapping of relational databases to RDF," W3C RDB2RDF Incubator Group Report, vol. 1, pp. 113-130, 2009.

[14] Spanos D.-E., Stavrou P., and Mitrou N., "Bringing relational databases into the semantic web: A survey," Semantic Web, vol. 3, pp. 169-209, 2012.

[15] Hazber M., Yang J., and Jin Q., "Towards integration rules of mapping from relational databases to semantic web ontology," in 2010 International Conference on Web Information Systems and Mining, 2010, pp. 335-339.

[16] Vavliakis K. N., Grollios T. K., and Mitkas P. A., "RDOTE–Publishing Relational Databases into the Semantic Web," Journal of Systems and Software, vol. 86, pp. 89-99, 2013.

[17] Sequeda J. F., Arenas M., and Miranker D. P., "On directly mapping relational databases to RDF and OWL," in Proceedings of the 21st international conference on World Wide Web, 2012, pp. 649-658.
[18] Bumans G., "Mapping between Relational Databases and OWL Ontologies: an example," Scientific Papers, University of Latvia, vol. 756, pp. 99-117, 2010.
[19] Jain V. and Prasad S., "Mapping Between RDBMS And Ontology: A Review," International Journal of Scientific & Technology Research, vol. 3, pp. 307-313, 2014.
[20] Čerāns K. and Būmans G., "RDB2OWL: a language and tool for database to ontology mapping," in Proceedings of the CAiSE 2015 Forum at the 27th International Conference on Advanced Information Systems Engineering (CAiSE 2015), Kista, Sweden, 2015, pp. 81-88.
[21] Hazber M. A., Li R., Zhang Y., and Xu G., "An approach for mapping relational database into ontology," in 2015 12th Web Information System and Application Conference (WISA), 2015, pp. 120-125.
[22] Jain V. and Singh M., "A framework to convert relational database to ontology for knowledge database in semantic web," International Journal of Scientific & Technology Research, vol. 2, pp. 9-12, 2013.
[23] Arenas M., Bertails A., Prud'hommeaux E., and Sequeda J., "A direct mapping of relational data to RDF," W3C recommendation, vol. 27, pp. 1-11, 2012.
[24] Sequeda J. F., Tirmizi S. H., Corcho O., and Miranker D. P., "Survey of directly mapping SQL databases to the Semantic Web," Knowledge Engineering Review, vol. 26, pp. 445-486, 2011.
[25] Souripriya D., Seema S., and Richard C. (2012). R2RML: RDB to RDF mapping language. W3C Recommendation, World Wide Web Consortium, September 2012. Available: http://www.w3.org/TR/r2rml/.
[26] Beckett D., Berners-Lee T., Prud'hommeaux E., and Carothers G., "RDF 1.1 Turtle," World Wide Web Consortium, 2014.
[27] Auer S., Dietzold S., Lehmann J., Hellmann S., and Aumueller D., "Triplify: light-weight linked data publication from relational databases," in Proceedings of the 18th international conference on World wide web, 2009, pp. 621-630.
[28] Bizer C. and Cyganiak R., "D2r server-publishing relational databases on the semantic web," in Poster at the 5th international semantic web conference, 2006.
[29] Barrasa J., Corcho Ó., and Gómez-Pérez A., "R2O, an extensible and semantically based database-to-ontology mapping language," in Proceedings of the 2nd Workshop on Semantic Web and Databases, Toronto, Canada, 2004.
[30] OpenLink Virtuoso Universal Server. Available at: http://virtuoso.openlinksw.com.
[31] DB2Triples. Available: https://github.com/antidot/db2triples.
[32] Morph-RDB. Available: https://github.com/oeg-upm/morph-rdb.
[33] RDF Graph Features | Oracle Spatial and Graph. Available: https://www.oracle.com/database/technologies/spatialandgraph/rdf-graph-features.html.
[34] RDF-RDB2RDF. Available: https://metacpan.org/release/RDF-RDB2RDF.
[35] XSPARQL. Available: https://sourceforge.net/projects/xsparql/files/xsparql/.
[36] Sequeda J. F. and Miranker D. P., "Ultrawrap: SPARQL execution on relational data," Journal of Web Semantics, vol. 22, pp. 19-39, 2013.
[37] Ultrawrap. Available: http://capsenta.com/ultrawrap.
[38] Boris Villazón-Terrazas M. H. (2012). RDB2RDF Implementation Report. Available: https://www.w3.org/TR/rdb2rdf-implementations/.
[39] Boris Villazón-Terrazas M. H. (2012). R2RML and Direct Mapping Test Cases. Available: https://www.w3.org/TR/rdb2rdf-test-cases/.
[40] Daoud A. M., Hosny K. M., and Mohamed E. R., "Semantic Representation of Music Database Using New Ontology-Based System," Journal of Theoretical and Applied Information Technology, vol. 98, 2020.
[41] Kumar A. P., Kumar A., and Kumar V. N., "A comprehensive comparative study of SPARQL and SQL," International Journal of Computer Science and Information Technologies, vol. 2, pp. 1706-1710, 2011.

[42] Alaoui L., Abatal A., Alaoui K., Bahaj M., and Cherti I., "SQL to SPARQL mapping for RDF querying based on a new Efficient Schema Conversion Technique," International Journal of Engineering Research and Technology, vol. 4, 2015.

[43] Soussi N. and Bahaj M., "Semantics preserving SQL-to-SPARQL query translation for Nested Right and Left Outer Join," Journal of applied research and technology, vol. 15, pp. 504-512, 2017.

[44] Rodriguez-Muro M., Kontchakov R., and Zakharyaschev M., "Ontology-based data access: Ontop of databases," in International Semantic Web Conference. Springer, Berlin, Heidelberg, 2013, pp. 558-573.

[45] Calvanese D., Cogrel B., Komla-Ebri S., Kontchakov R., Lanti D., Rezk M., et al., "Ontop: Answering SPARQL queries over relational databases," Semantic Web, vol. 8, pp. 471-487, 2017.

[46] Rodrıguez-Muro M., Hardi J., and Calvanese D., "Quest: efficient SPARQL-to-SQL for RDF and OWL," in 11th International Semantic Web Conference ISWC, 2012, p. 53.

[47] Rodriguez-Muro M., Rezk M., Hardi J., Slusnys M., Bagosi T., and Calvanese D., "Evaluating SPARQL-to-SQL translation in ontop," in In Proc. of the 2nd Int. Workshop on OWL Reasoner Evaluation (ORE), volume 1015 of CEUR Electronic Workshop Proceedings, pages 94– 100, 2013.

[48] Kontchakov R., Rezk M., Rodriguez-Muro M., Xiao G., and Zakharyaschev M., "Answering SPARQL queries over databases under OWL 2 QL entailment regime," in International Semantic Web Conference, 2014, pp. 552-567.

[49] Rodriguez-Muro M. and Rezk M., "Efficient SPARQL-to-SQL with R2RML mappings," Journal of Web Semantics, vol. 33, pp. 141-169, 2015.

[50] Brickley D. and Miller L. (2007). FOAF vocabulary specification 0.91. Available at: http://xmlns.com/foaf/spec/.

[51] Golbeck J. and Rothstein M., "Linking Social Networks on the Web with FOAF: A Semantic Web Case Study," in AAAI, 2008, pp. 1138-1143.

[52] Mika P., "Flink: Semantic web technology for the extraction and analysis of social networks," Journal of web semantics, vol. 3, pp. 211-223, 2005.

[53] Tang J., Zhang D., and Yao L., "Social network extraction of academic researchers," in Seventh IEEE International Conference on Data Mining (ICDM 2007), 2007, pp. 292-301.

[54] Nasution M. K. and Noah S. A., "Extraction of academic social network from online database," in 2011 International Conference on Semantic Technology and Information Retrieval, 2011, pp. 64-69.

[55] Sander M., Waltinger U., Roshchin M., and Runkler T., "Ontology-based translation of natural language queries to SPARQL," in 2014 AAAI Fall Symposium Series, 2014.