An Exploratory Study of RDF: A Data Model for Cloud Computing

A. Clara Kanmani, T. Chockalingam and N. Guruprasad

Abstract Semantic web is an extension of the web which focuses on the meaning of data content rather than the structure of the data content. It promotes many standards by the world wide web consortium (W3C). RDF is a data interchange standard widely used by semantic web community. Cloud computing is a computing paradigm which involves outsourcing of computing resources with the capabilities of resource scalability, on demand provisioning with little or no up-front IT infrastructure investment costs. Resource Description framework is a semantic data model for cloud computing. This paper analyze RDF in terms of its present status, comparison between RDF and traditional data models, its usage in semantic web data management, overview of semantic web rule languages and finally its limitations in representing concepts are examined.

Keywords Semantic web · Cloud computing · RDF · Data model

1 Introduction

Semantic web hush-up the vision of web enriched by annotations which allows both software and humans to use the content by making the web to some degree machine comprehensible. One of the key technologies used is resource description framework. It is a basis for dealing out XML based metadata which provides interop-

A.C. Kanmani (✉) · T. Chockalingam
Department of Computer Science and Engineering, Visvesvaraya Technological University, Research Resource Center, Belagavi 590018, India
e-mail: claracalton.joseph@gmail.com

T. Chockalingam
e-mail: tchocks@gmail.com

N. Guruprasad
Department of Computer Science and Engineering, New Horizon College of Engineering, Bengaluru 560103, India
e-mail: nguruprasad18@gmail.com

S.C. Satapathy et al. (eds.), *Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications*, Advances in Intelligent Systems and Computing 516, DOI 10.1007/978-981-10-3156-4_68

erability between applications that exchange machine understandable information on the web.

Semantic Web technologies has wide range of applications in data integration, resource discovery and classification, which provides a chance to integrate the data in different locations and to perform an efficient search. It further can be used in cataloging, which describes the content and content relationships, available at a particular web site and knowledge sharing and exchange.

1.1 Architecture of the Semantic Web Stack

The semantic web stack [1], encompasses three layers, as depicted in Fig. 1.

Lower layers includes IRI, Internationalized Resource Identifier (IRI), which, identifies semantic web resources and Unicode serves to represent and manipulate text in many languages. XML is a markup language that makes possible creating documents consisting of structured data.

Middle layers includes RDF, a graph data model which represents web resources as subject, object and predicate, together known as triples. RDF vocabulary is well described in RDFS. Advanced constructs to express RDF statements are available in web ontology language (OWL). Users who want to query any kind of RDF data can utilize the RDF query language SPARQL. There is standardization required in the upper layers.

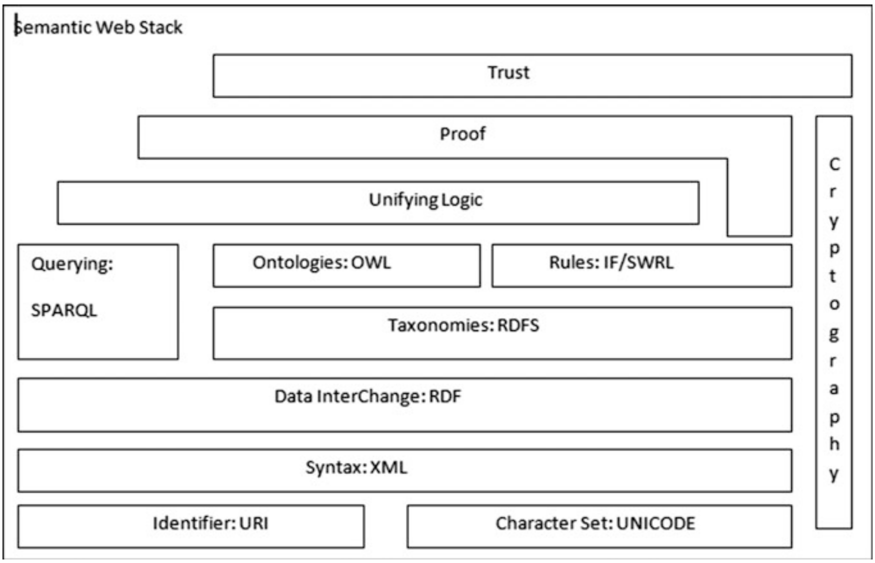

Fig. 1 Semantic web stack

1.2 Semantic Web and Cloud Computing

Cloud computing is a budding paradigm in data processing communities. Businesses store and access data at remote locations in the cloud. Data is the most vital asset the application generates over time and is essential for continued functioning. As the popularity of cloud computing grows, there are many challenges to be addressed by service provider. They have to maintain huge quantities of heterogeneous data while providing efficient information retrieval.

The effect of an inefficient data model may reduce an organization business performance on cloud environment. The change might even required to carry across different data models. At the same time semantic web is also an emerging area to augment human reasoning. A competent data model for cloud computing can be constructed using RDF.

This paper is organized as follows. Section 2 gives an introduction to RDF, design goals and evolution, Sect. 3 makes an attempt to perform an analysis taking into account of some of the vital criteria pertaining to it and finally Sect. 4 ends up with analysis and discussion followed by conclusion.

2 Resource Description Framework

Resource Description framework is an infrastructure that enables the encoding, exchange and reuse of structured metadata. This session gives a brief of evolution of RDF, introduction to RDF and design goals of RDF.

2.1 Evolution of RDF

Content rating has information about the content of web pages. It also gives an idea about whether this information is given by a qualified researcher. In this context, PICS (Platform for Internet Content Selection) is a metadata framework which provides content rating. RDF as a general metadata framework and knowledge representation mechanism is the inspiration obtained from PICS.

2.2 Introduction to RDF

The Resource Description Framework (RDF) is a framework for expressing information about resources. Resources can be anything, including documents, people, physical objects, and abstract concepts, as shown in Fig. 2.

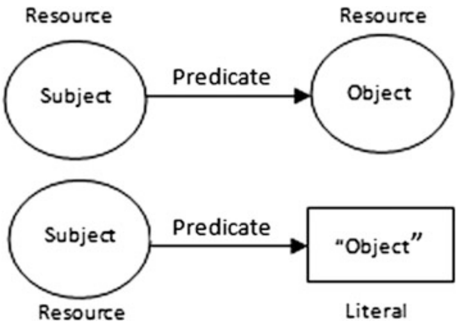

Fig. 2 RDF graph

It is a simple language description model that makes statements about a resource by either defining its relationships with other resources or by defining its attribute. The resource is the subject, the relationship is the predicate, and the attribute value is the object.

2.3 Design Goals of RDF

The design of RDF is proposed to meet the following goals, to have a simple model, formal semantics and provable reference. It focuses on using an extensible URI-based vocabulary and XML-based syntax. It supports the use of XML schema data types and allows anyone to make statements about any resource.

The aim of RDF is to have a mechanism for expressing resources where no assumptions can be made about a particular application domain, nor the semantics of any domain is defined.

3 RDF: An Analysis

An analysis about RDF is carried in different ways. To begin with Sect. 3.1 makes a comparison between RDF and traditional data model to verify whether RDF becomes a unique ontological framework for conceptual representation compared to an ER model or an Object oriented data model. Section 3.2 look at the present status of RDF as accepted by semantic web community. Section 3.3 deals with RDF usage in semantic web data management. Finally, Sect. 3.4 examines the limitations of current RDF in representing concepts followed by analysis and discussion and finally ends up with conclusion.

3.1 Comparison Between RDF and Traditional Data Models

RDF is a model of entities and relationships [2]. In ER model there are entity types, and there is a set of relationships defined for it. The RDF model is the same as ER model but the relationships are of prime importance, they are identified by a URI. ER model is a static data model representing raw data, entities and relationships. Raw data is inferred. ER model is suited for structured data modeling. It becomes a messy when dealing with semi-structured content.

The RDF model is a dynamic data model. It captures the perceived semantics and dynamically relations can be added in RDF. Any web resource is expressed in the form subject, predicate and object called triples The RDF model does not have methods and all parts of the RDF graph are public.

There is a direct connection between a model of relational databases with semantic data model [2]. Any relational database has tables, which possess rows, or records. Each record has a set of fields. Any record is a RDF node, column name denotes a predicate or property and record field is object which can be literal.

RDF data is managed by Relational database systems, in a different way. Though the relational data model is the most prevalent one, it has become unsuitable for data federation purpose.

At whatever point any sort of new information with various attributes arrive, it can undoubtedly be attached to the current model without changing the current schema. This flexibility nature of RDF could be seen as an information driven configuration. This very flexibility nature of RDF makes one to view it as information driven outline. This information driven configuration makes RDF an exceptional ontological system contrasted with other conventional information models. Figures 3 and 4 shows the dynamic nature of RDF about how relations can be added in RDF.

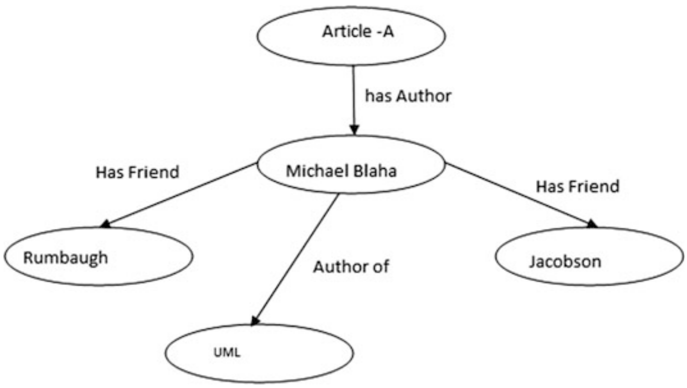

Fig. 3 RDF schema

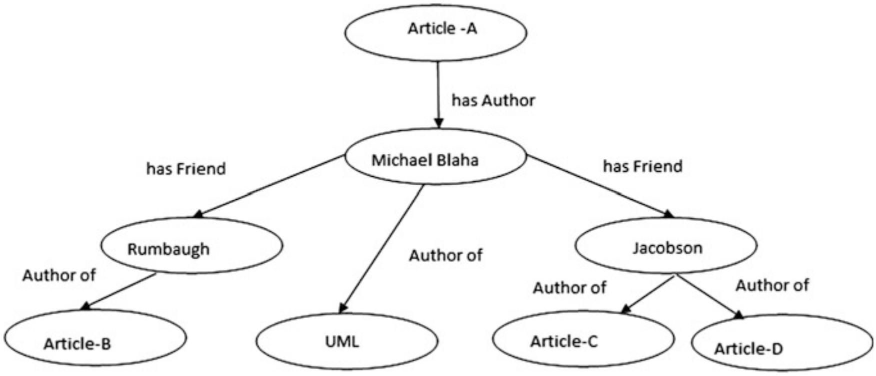

Fig. 4 Modified RDF schema

3.2 RDF Present Status

There is an RDF interest group [3] who provides a forum for discussion of a wide range of RDF related activities and issues. Another adaptation of RDF may incorporate changes regarding components, semantics, and serialization punctuations, in reverse similarity is of a central significance. In fact, RDF has been conveyed by apparatuses and applications, and the most recent couple of years have seen a noteworthy uptake of Semantic Web innovations and production of billions of triples coming from open databases (e.g. the Linked Open Data people group).

It would be, accordingly, disadvantageous to this advancement if RDF was seen as shaky and if the legitimacy of current application would be made due by a future development. As a result, with any progressions of RDF [4], in reverse similarity prerequisites ought to be formalized along the accompanying bearings, any substantial RDF graphs as far as the RDF 2004 adaptation ought to stay legitimate regarding another rendition of RDF, change or changes on the RDF (S) semantics.

There is a prerequisite for clarifications, blunders, updates on the RDF (S) specifications. There can be an upgrade in the zone of serialization formats, center RDF (S) highlights, vocabulary terms those applications require. Different ranges of advancement incorporates following changes in RDF graphs, how provenance fits in RDF and tending to best practices (use of rdf: value, reification).

3.3 Semantic Web Rule Languages

Haase et al. [5] makes a comparison of semantic web query languages. Harold Boley et al. gives an overview [6] of RuleML, a rule markup language for the semantic web. Modular syntax and semantics of RuleML and current RuleML 0.8 DTDs are presented. Negation handling, priorities and implementations of RuleML via XSLT is been analysed. Boley et al. [6] also provides the requirements for further RuleML versions.

Horrocks et al. (2004) puts forth a proposal [7] for SWRL: A Semantic Web Rule Language which combines OWL and ruleML. SWRL incorporates an abnormal state dynamic sentence structure for Horn-like guidelines in both the OWL DL and OWL Lite sublanguages of OWL. A XML language structure taking into account RuleML and the OWL XML Presentation Syntax and also a RDF solid linguistic structure in light of the OWL RDF/XML trade grammar are likewise determined, with a few illustrations.

3.4 RDF Usage in Web Data Management

The volume of RDF information keeps on becoming over the previous decade and numerous known RDF datasets have billions of triples. Expansive measure of semantic information are accessible in the RDF group in numerous fields of science, designing, and business, including bioinformatics, life sciences, business knowledge and interpersonal organizations. An incredible test of dealing with this tremendous RDF information is the means by which to get to this huge RDF information productively.

In this [8] a well known methodology is been proposed for tending to the issue which constructs a full set of permutations of (S, P, O) records. Despite the fact that this methodology has appeared to quicken joins by requests of size, the expansive space overhead constrains the adaptability of this methodology and makes it heavyweight. TripleBit [8] a quick and reduced framework for putting away and getting to RDF information is been planned.

A structure list for RDF is anticipated [9] which can be utilized for questioning RDF data for which the schema is fragmented or not accessible. In light of data caught by the structure file, comparably organized information components are physically assembled and put away adjacently on disk.

At questioning time, the list is utilized for “structure-level” handling to distinguish the gatherings of information that match the inquiry structure. Structure-level preparing is then consolidated with standard “data level” operations that include recovery and join systems executed against the information.

Proposed methodology mentioned in this [9] indicates adaptability results, i.e., execution change builds more than straightly with the span of the information. This is on the grounds that its execution does not entirely correspond with information measure but rather relies on upon the heterogeneity of structure examples displayed by the information.

3.5 Limitations of RDF

New serializations groups (e.g. Turtle) have picked up a huge backing by the group, while the entanglements in RDF/XML punctuation have made a few challenges in practice [10] and also in the acknowledgment of RDF by a bigger Web people group. At last, at present there is no standard programming API to oversee RDF information, the necessity may emerge to characterize such a standard either in a general, programming dialect autonomous path or for a portion of the essential dialects (Javascript/ECMAscript, Java, Python, …).

There are a few confinements of current RDF diagram as for expressiveness of specific properties. These properties are depicted in detail [11] underneath.

- Local extent of properties

rdfs: range characterizes the scope of a property (e.g. eats) for all classes. In RDF Schema we can't announce range limitations that apply to a few classes only. E.g. we can't say that cows eat just plants, while different creatures may eat meat, as well.

- Disjointness of classes

Now and then we wish to say that classes are disjoint (e.g. tyke and adult).This property is hard to express in RDF.

- Boolean blends of classes

Now and then we wish to assemble new classes by consolidating different classes utilizing union, crossing point, and supplement. E.g. human is the disjoint of the classes youngster and grown-up.

4 Analysis and Discussion

RDF that begins as a simple graph data model further gets expanded into complex structures. It has more advantages compared to the conventional data models. Using RDF one can extend and adapt data and their organizational schema flexibly as per user requirements. There is a provision for exploration and analysis with this model which is not easily possible with other existing models. This model can be used without disturbing the pre-existing schema. Relational database has proved itself as a good database with structured data. It stores data in tables which saves space, ideal for the enterprises etc.

5 Conclusion

RDF is new and still emerging. There are lot of issues which are not addressed in RDF. Once those issues are identified, it could be represented with ease which is very much intricate in the in present scenario. Looking at the present status of RDF by semantic web community, lots of development is still to be done in this area and the main goal is to reuse, as much as possible, existing data in its existing form minimizing the RDF data which has to be created manually.

References

1. Semantic Web Stack, https://en.wikipedia.org/wiki/Semantic_Web_Stack
2. W3C Semantic Web Activity, http://www.w3.org/2001/sw/
3. RDF Interest Group, https://www.w3.org/RDF/Interest/
4. W3C Workshop, https://www.w3.org/2009/12/rdf-ws/cfp
5. P. Haase, J. Broekstra, A. Eberhart, R. Volz, A comparison of RDF query languages, in *ISWC 2004 LNCS* 3298 (2004), pp. 502–517
6. H. Boley, S. Tabet, G. Wagner, Design rationale of RuleML, in *A Markup Language For Semantic Web Rules* (2001)
7. I. Horrocks, P.F. Patel-Schneider, F. Van Harmelen, A semantic web rule language combining OWL and RuleML, in *W3C submission*
8. TripleBit: A Fast and Compact System for Large Scale RDF Data, in *Proceedings of VLDB*, vol. 6(7) 2013
9. T. Tran, G. Ladwig, S. Rudolph, Managing structured and semi structured RDF data using structure Indexes
10. http://answers.semanticweb.com/questions/11557/what-are-the-most-severe-limitations-of-rdfxml
11. http://www2.cs.man.ac.uk/~raym8/comp38212/main/node185.html