JSON and its use in Semantic Web

Mrinal Pandey
AIIT
Amity Institute of
Information Technology
Lucknow, India

Rajiv Pandey
AIIT
Amity Institute of
Information Technology
Lucknow, India

ABSTRACT

The semantic web has evolved over the current web and aims to provide a web that allows for easy retrieval and accessing of information by both man and machine. It provides for a wide variety of technology stacks , language standards and software components which help both man and machine to access data easily. Intelligent information retrieval and the credibility of data is managed in semantic web by the use of Provenance. However in addition to embedding provenance data there is also a need to facilitate its easy availability for the client side for the purpose of manipulation and transportation accross. This task is achieved by incorporating the use of JSON in Provenance thereby creating PROV-JSON serializations of the PROV-DM i.e the PROV data model. In this paper we have made an endeavor to create an ontology and have embedded provenance information on the client side using PROV-JSON instances.

General Terms

www, Intelligent web and Semantic web

Keywords

Semantic web , JSON, Linked data, Ontology,Provenance, trust

1. INTRODUCTION

The Semantic Web suggested by Tim Berners-Lee is an extension of the current World Wide Web.It aims at a web that provides for   information and services that are understandable and usable   by both human beings and by computers. Semantic Web is aims to resolve the following key limitations of the current Web [1]-[3]:

- Lack of structured information – a large no of documents present on the web lack structure;

- Ambiguity of web content - the poor aggregation of information to specific contexts;

- Inadequate mechnisms to automatically transfer information;

- Difficulty in maintaining and preserving   and providing trustworthy data on the web;

- Paucity of universal mechanisms to provide the capability to machines to understand and retrieve the desired information. [4]

 The Semantic Web allows for the definition and linking of information across heterogeneous domains so as to allow for the manipulation of information by   machines   for automation, integration and reuse across heterogeneous applications.   The Semantic intends   to make computers understand the human language, by   defining a universal model for the representation of information along with the constructs that enable these machines to easily understand the

information [5]. The Semantic Web intends to go beyond the current web and aims at providing mechanisms that hardwired computers and other devices to share contents across a wide variety of heterogeneous platforms[18]. The key features   of Semantic Web are as follows[3]:

- Extends the current web by, adding to it a concept layer;

- Provides for machine-interpretable   and   editable web content;

- Provides   for ways to semantically annotate current web documents thereby allowing for their easy organization, search and retrieval;

- Enhance the   search mechanisms by using Ontologies.

- Empowers the software agents to automatically carry out sophisticated tasks using smart data.

- Nurture enhanced communication between platform-independent software agents;

- Provides trustworthy data to a seeker of information by the use of Provenance.

2. ONTOLOGY

Ontology is a   description of all concepts or meanings as well as their relationships amongst these areas of knowledge such as Medicine, Mathematics, etc. To provide a wide variety of representation for a wide diversity of concepts from the real world. Ontologies allow for easy extension, refinement and reuse by other ontologies[19]. Though Ontologies are able to successfully represent and retrieve semantic data by means of annotations yet there is a need to incorporate the element of trust in semantic documents. The element of trust can be incorporated by using provenance in the semantic web. Provenance deals with the origin or the source of a piece of data item which helps the user in ascertaining its credibility. However In addition to incorporating   provenance the provenance assertions should be built catering to their easy availability, accessibility and transportation over the web and the software agents or the clients. Contemporary research focuses on using XML as a convenient and platform independent way for transmitting data over the web to the client side. In our paper we have thus made an endeavor to build client side representations of the PROV-DM instances using a new data interchange format namely-JSON or the Java script object notation. The paper further thus describes Provenace, JSON, Relatiiship of JSON with semantic web and Linked data and finally the creation of JSON instances of the PROV-DM data model.

3. PROVENANCE

The general understanding of provenance is the source or history of a derivation of a particular object. Provenance holds an important place in many practical fields. For example, the Food and Drug Administration department can use

provenance to keep a record of the drug discovery as long as the drug is in use. In aerospace engineering, simulation records are required to kept for a period of 99 years which provide for important provenance information. In museum and archive management provenance helps in maintaining the archival history regarding the acquisition of a record[12]. Provenance thus is crucial in ensuring the trust ability of documents across various domains. In our paper however we have tried to Implemented in the provenance in a University peoples program Ontology and have implemented provenance into the same using JSON serialization to facilitate its easy availability, manipulation and transportation across client side.

4. JSON

JSON (Javascript Object Notation) [6] is a language-independent data-interchange format that has been derived from the ECMAScript (JavaScript). JSON objects are converted to string arrays, permit higher parsing efficiency and can be prepared and transported easily as compared to other heavier counterparts like XML[7]. It uses the conventions that are used by languages like c, c++, java etc which makes it an ideal language for data inercahnge. JSON objects comprise of primarily two data structures: a collection of name-value pairs and an ordered list of values[7].Unlike XML JSON uses closed tags and its objects are analysed as strings instead of DOM which makes it easy to parse and generate codes written in JSON[8]. Moreover, [9] states clearly that all the features of XML are fulfilled by JSON[10].A simple example of JSON name value pair is demonstrated below-[8]

JSON object structure.

. { "contacts":

 { "contact":

{ "name": "Jack Sparrow",

"phone": "123-532-5392"

"work": "actor"

}

 }

 }

The above can also be represented in the form of a Graph as illustrated in the figure provided below-[8]
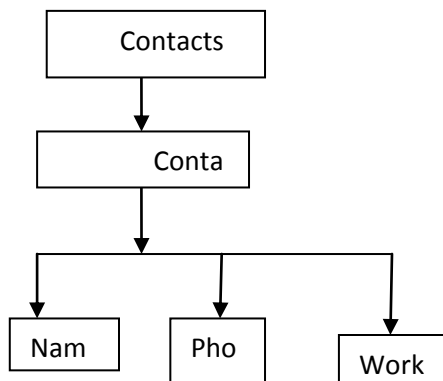
5. JSON AND ITS USE IN SEMANTIC WEB AND LINKED DATA-

Semantic web also known as the intelligent web is not only about putting data on the web it also aims about linking the data to enable both man and machine to explore the web of data thereby formulating the concept of Linked Data. Linked data, a term first proposed by Tim Berners Lee is a term that is used to describe data relationships between interconnected documents. Some of the basic features of Linked data include-

- Use of URI's for naming things.

- Using HTTP URI's for enabling people to easily access those names.

- Providing useful information by means of the URI's.

- Enabling people to discover more things by providing links to other URI's[13].

Semantic web and Linked data applications often use the concept of triples to represent documents on the web. Often these triples are also represented as graphs. JSON serves as a suitable data format by Semantic web and linked data applications in such cases. This comes from the simplicity that JSON possesses. JSON provides a standard way to represent data . The new form of JSON i.e JSON-LD allows for the addition of context to data and produces coerced values which at their very basic form look like simple JSON structures and produce fully linked data graphs that can be processed very easily.[11] Provenance serves as a crucial component of semantic web because it helps in incorporating the element of trust in the semantic web application. However modern day applications in addition to requiring provenance also require that the provenance information generated by the semantic web applications is platform independent, transportable and easy to generate and manipulate. This can be effectively achieved if use the JSON instead of the traditional XML to generate provenance. This can be achieved by the use PROV-JSON serialization at the client end instead of the conventional PROV-N, PROV-O and PROV-XML serializations. In the following section we have thus taken an Ontology and have illustrated the creation of PROV-JSON instances for entities, events and activities which can thus be produced and can be easily processed using the stated serialization.

6. PROVENANCE IN UNIVERSITY PEOPLE PROGRAM ONTOLOGY-

Ontologies are"vocabularies of representational terms, classes, relations, functions, object constants with agreed upon definitions ,in the form of human readable text and machine-enforceable, declarative constraints on their well-formed use".[6] Ontology are only deemed to be effective if it has trust associated with it.

Trust , a crucial component can be incorporated by embedding provenance inside the Ontology. In This section we have considered the example of University People Program Ontology constructed using OWL and have described the procedure to embed provenance information into it.

"Below can be found a code of the University People program Ontology which allows us to describe the-three basic components i.e. Entity, Activity and Agent into which an

Ontology is mapped for embedding Provenance information into it.

A snippet of the University People Program is as follows-[14]"

```
<ObjectPropertyAssertion>
 <ObjectProperty IRI="#teaches"/>
         <NamedIndividual IRI="#Rajiv"/>
          <NamedIndividual IRI="#mobileComputing"/>
     </ObjectPropertyAssertion>
     <ObjectPropertyAssertion>
     <ObjectProperty IRI="#isTaughtBy"/>
     <NamedIndividual IRI="#mobileComputing"/>
     <NamedIndividual IRI="#Rajiv"/>
 </ObjectPropertyAssertion>
```

The above code states that Rajiv who is a professsor is linked to mobile computing, which is a subject through object property teaches & vice versa through object property 'isTaughtBy'.[14]

The above OWL code is mapped to Entity, Agent and activity which are the crucial components of the PROV-DM model. Rajiv who is a professor acts as an Agent that performs the activity of Compiling and Authoring, thereby generating the Entity called MobileComputing TextBook. The activity of compiling & authoring is performed by compilingTopicFrom various research papers (which are themselves entities). However other research papers that result from the given research paper may be considered as entities and the given research paper as an agent, as the given research paper contributes towards the development of the other research papers. On adding PROV descriptions is \to the manipulating agents the resultant Entity MobileComputing Text Book is known to be a trusted document[14][24]

```
<ObjectPropertyAssertion>
 <ObjectProperty IRI="#teaches"/>
         <NamedIndividual IRI="#Rajiv"/>
          <NamedIndividual IRI="#mobileComputing"/>
     </ObjectPropertyAssertion>
     <ObjectPropertyAssertion>
     <ObjectProperty IRI="#isTaughtBy"/>
     <NamedIndividual IRI="#mobileComputing"/>
     <NamedIndividual IRI="#Rajiv"/>
 </ObjectPropertyAssertion>
```

The above code describes that individual Rajiv is linked to individual mobile computing through object property teaches & vice versa through object property 'isTaughtBy'.[14]

The above OWL code is mapped to provenance components of Entity, Agent and activity . Rajiv a professor ,is the Agent who performs the activity of Compiling and Authoring, thereby generating the Entity i.e., MobileComputing TextBook along with Mrinal , another agent and acts as a contributor to the activity while performing the above activity.

Once the book is compiled it is published by the World Wide Web consortium and an updated version of the same is generated by the consortium.

Below we show the creation of PROV-JSON serialization of PROV-N instances by identifying the entity, agent and activities such as derivation, usage, and generation activities involved in a serialization. Further we also show the different views of the provenance with respect to t6he creator of the provenance i.e the author , the manipulation of provenance information by the processes that are using it and then finally how the provenance can be manipulated when present in various bundles.

Authors view-[15][16][17]

```
{
"prefix": {

    "ex": "http://example.org",
    "w3": "http://www.w3.org/",
    "tr": "http://www.w3.org/TR/2011/"
  },
"entity":{
"tr:wd-doc1":{
"ex:name":"MCT",
"prov:type":"document"
            },
"tr:wd-rp1":{
"ex:name":"rep1",
"prov:type":"document"
}


},
"agent":
{
"ex:ag1":{
"ex:name":"rajiv pandey",
"prov:type":{
"$":"prov:person",
"prov:type":"xsd:QName"}    }
"ex:ag2":{"ex:name":"Mrinal",
"prov:type":          {
"$":"prov:person",
"type":"xsd:QName"
                }
        }
},
"activity":
```

{

"ex:ac1":

{

"prov:type":"creation and internal editing"

}

},

"wasAssociatedWith":{

"_:ASS1":

{

"prov:activity":"ex:ac1",

"prov:agent":"ex:ag1",

"prov:role":"editor"

},

"_:ASS2":

{

"prov:activity":"ex:ac1",

"prov:agent":"ex:ag2",

"prov:role":"contributor"

},

"_:ASS3":

{

"prov:activity":"ex:ac1",

"prov:agent":"ex:ag1",

"prov:role":"compiler"

}

},

"wasGeneratedBy":{

"_:GN1":

{

"prov:activity":"ex:ac1",

"prov:entity":"tr:wd-doc1" }},

"wasDerivedFrom":{

"_Drv1":{

"prov:generatedEntity":"tr-wd-doc1",

"prov:usedEntity":"rp1"

}}}

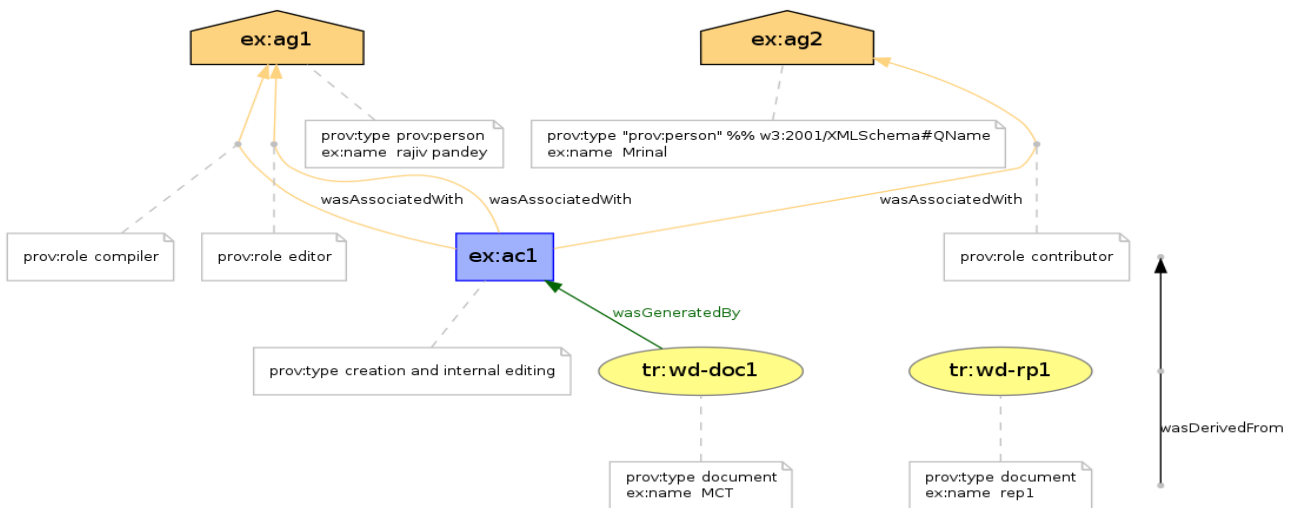**Fig1.Diagram depicting the relationship between Entities, agents and activities in authors view[15][16]**

Document

 prefix w3 <http://www.w3.org/>

 prefix tr <http://www.w3.org/TR/2011/>

 prefix ex <http://example.org>

 wasGeneratedBy(tr:wd-doc1, ex:ac1, -)

 activity(ex:ac1, -, -
, [prov:type="creation and internal editing"])

 entity(tr:wd-
doc1, [prov:type="document", ex:name="MCT"])

 entity(tr:wd-rp1, [prov:type="document", ex:name="rep1"])

 agent(ex:ag1, [prov:type="prov:person", ex:name="rajiv pan
dey"])

 agent(ex:ag2, [prov:type="prov:person" %% w3:2001/XML
Schema#QName, ex:name="Mrinal"])

 wasDerivedFrom(-, -, -, -, -)

 wasAssociatedWith(ex:ac1, ex:ag1, -
, [prov:role="compiler"])

 wasAssociatedWith(ex:ac1, ex:ag2, -
, [prov:role="contributor"])

 wasAssociatedWith(ex:ac1, ex:ag1, -, [prov:role="editor"])

end document

Prov-N representation of the corresponding Author view

Fig 2-Diagram showing the document metrics for authors view created using Prov-Store[15][16][17]

Process view-[15][16]

```
{
"prefix": {
    "ex": "http://example.org",
    "process": "http://www.w3.org/2005/10/Process-
20051014/tr.html#",
"rec54": "http://www.w3.org/2001/02pd/rec54#",
    "tr": "http://www.w3.org/TR/2011/",
    "w3": "http://www.w3.org/",
    "email": "https://lists.w3.org/Archives/Member/w3c-
archive/"
    },
"entity":{
"tr:wd-doc2":{                                    // here doc2
is analogous to tr:WD-prov-dm-20015
            "prov:type":{
"$":"rec54:WD",
"type":"xsd:QName"
            }
},
"activity":{
"ex:act1":{
            "prov:type":"publish"
            }
},
"wasGeneratedBy": {
    "_:GN2": {
        "prov:activity": "ex:act1",
```

```
        "prov:entity": "tr:wd-doc2"          // here doc2
is analogous to tr:WD-prov-dm-20015
    }
},
"wasDerrivedFrom":{
        "_:D1":{
"prov:generatedEntity":"tr:wd-doc2",      // here doc2
is analogous to tr:WD-prov-dm-20015
"prov:usedEntity":"tr:wd-doc1"   }        // here doc2
is analogous to tr:WD-prov-dm-20018
},
"used":{
"_:u1":{
    "prov:activity":"ex:act1",
    "prov:entity": "email:2011Dec/0111"
        }},
"wasAssociatedWith": {
    "_:ASS5": {
        "prov:activity": "ex:act1",
        "prov:agent": "w3:Consortium",
        "prov:plan": "process:rec-advance"
    }
  }
}
```
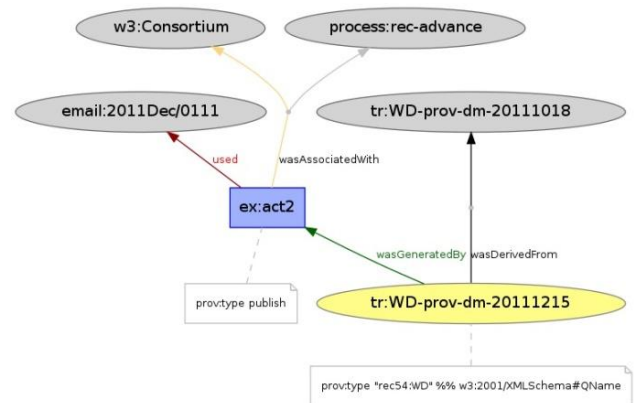

Diagram depicting relationships between entities in process view[15]

Provenance attribution-[15]

Provenance attribution can be specified by means of creating a Bundle for the same. A bundle is a data structure that allows a provenance definition created for one entity to be specified for another entity. This is advantageous when we wish to relate provenance generated by two different parties. It provenance trustworthiness to documents created by users by providing for provenance of provenance. The idea states that for the users to trust the provenance of a document the source that generates the provenance must also be trustworthy. In the following illustration we thus provide for the linking of the provenance created from two different prespectives by two

different parties by means of Bundles thus depicting how the two documents are related to each other.

Linking the two perspectives of Provenance by means of Bundles-[15][16]

```
          {
  "prefix": {

    "ex": "http://example.org",
    "w3": "http://www.w3.org/",
  },
  "bundle": {

    "ex:author-view": {
      "agent": {

        "ex:rajiv pandey": {
          "prov:type": {
            "$": "prov:Person",
            "type": "xsd:QName"
          }
        },

        "ex:Mrinal": {
          "prov:type": {
            "$": "prov:Person",
            "type": "xsd:QName"
          }
        }
      },

    },

    "ex:process-view": {
      "agent": {

        "w3:Consortium": {
          "prov:type": {
            "$": "prov:Organization",
            "type": "xsd:QName"
          }
        },

        }
```

```
      },

    }
}
  ,
    "entity": {

      "ex:author-view": {
        "prov:type": {
          "$": "prov:Bundle",
          "type": "xsd:QName"
        }
      },

      "ex:process-view": {
        "prov:type": {
          "$": "prov:Bundle",
          "type": "xsd:QName"
        }
      }
    },
    "wasAttributedTo": {

      "_:wAT7": {
        "prov:agent": "ex:rajiv",
        "prov:entity": "ex:author-view"
      },
      "_:wAT9": {
        "prov:agent": "ex:mrinal",
        "prov:entity": "ex:author-view"
      },
      "_:wAT8": {
        "prov:agent": "w3:Consortium",
        "prov:entity": "ex:process-view"
      }
    }}
```

It is thus demonstrated that provenance when added to OWL can ensure that the trust ability of the semantic document which is attained because the software agents or machines can now

The above stated provenance descriptions can be effectively added to semantic web pages to make them easily accessible to the software clients . Thus by combining the power by combining the power and simplicity of JSON with that of provenance the following advantages can be easily achieved-

- Interoperable: Allows for the easy conversion of instances written in XML, OWL and PROV Notations to be converted to and from PROV-JSON

- Lightweight: makes use of the constraints and data types provided in JavaScript.

- Ease of use: easy to use and create because of the presence of inbuilt namespaces.

- Scalable: allows for the creation of custom tags in JSON.

7. CONCLUSION

Considering the fact that provenance information needs to be shared across vast domains and various types of clients use of PROV-JSON as a common medium for understanding and Interchange becomes all the more important. A Description of PROV-JSON has been provided in this paper. The description clearly highlights the significance of agent expression, entity expression and the activities expressions created using JSON. The paper thus clearly highlights the relevance and power of JSON in creating Provenance assertions with respect to different clients.

8. REFERENCES

[1] Berners-Lee, T., Hendler, J., Lassila, O,2001. The Semantic Web. Scientific American (2001). URL: http://www.sciam.com/article.cfm?articleID=000481441 0D2-1C70-84A9809EC588EF21\&sc=I100322.

[2] Lu, S., Dong, M., Fotouhi, F, 2002.The Semantic Web: opportunities and challenges for next-generation Web applications. Information Research, Vol. 7 No. 4, 2002. URL: http://informationr.net/ir/7-4/paper134.html.

[3] Pereira R.G. and Freire M.M., 2005. "Semantic Web", in Encyclopedia of Multimedia Technology and Networking, M. Pagani (ed.), Idea Group, Inc., 2005.

[4] Pereira Rui. G.and. Freire, Mário M,2006.SWedt: A Semantic Web Editor Integrating Ontologies and Semantic Annotations with Resource Description Framework. Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW 2006)

[5] Berners-Lee,T.., (1998)."What the Semantic Web can represent", W3C, 1998. URL: http://www.w3.org/DesignIssues/RDFnot.html

[6] Online :Introducing JSON. http://www.json.org/.

[7] Abd El-Aziz,A.A. , Kannan,A. ,2014.JSON encryption-2014 International Conference on Computer Communication and Informatics (ICCCI -2014), Jan. 03 – 05, 2014, Coimbatore, INDIA

[8] Wang Guanhua,2011Improving Data Transmission in Web Applications via the Translation between XML and JSON-2011 Third International Conference on Communications and Mobile Computing

[9] Wang,G.,2011. Improving Data Transmission in Web Applications via the Translation between XML and JSON. In Proceedings of 2011 Third International Conference on Communications and Mobile Computing (CMC), pages 182 – 185, 18-20 April, 2011.

[10] Griffin,L., Butler ,B., E. de Leastar, B. Jennings, and D. Botvich,2012. On the Performance of Access Control Policy Evaluation. In Proceedings of 2012 IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY), pages 25 – 32, 16-18 July, 2012.

[11] Berners-Lee,T.,. World Wide Web Consortium, 18 June 2009. JSON LD primer-W3C Note. URL: http://www.w3.org/DesignIssues/LinkedData

[12] Chen Liming , Yang Xueqiang, Tao Feng,2006.A Semantic Web Service Based Approach for Augmented Provenance. Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06).

[13] Online:URL: http://www.w3.org/DesignIssues/LinkedData

[14] Rajiv Himanshu,2014. "Evaluating the need of PROV in OWL", ACEEE, IJRTET, 2014

[15] Online:PROV-STORE-https://provenance.ecs.soton.ac.uk/store/account/login/

[16] Online:https://provenance.ecs.soton.ac.uk/store/documents/114345/

[17] Online:https://provenance.ecs.soton.ac.uk/store/documents/115642/

[18] Haibo,Yu, Tsunenori, Mine, Makoto Amamiya,2012 "Balance: A key factor for the evaluation of semantic web applications"Proceedings of IIAI International Conference on Advanced Applied Informatics,2012

[19] .SaleenaJ,.B, Salini , M.G., Venkateswaran, Siva,2010.A Semantic Approach to Construct a Knowledge Portal for E-Learning Using Ontology".Proceedings of 4th International Conference on Distance Learning and Education (ICDLE)(2010).

[20] Online:Protégé tool. URL: http://protege.stanford.edu/.