AUTHORIZATION FRAMEWORK FOR MEDICAL DATA

Geetha Madadevaiah[1] , RV Prasad[1], Amogh Hiremath[1], Michel Dumontier[2], Andre Dekker[3]

[1]Philips Research, Philips Innovation Campus, Philips India Ltd, Manyata Tech Park, Bangalore
[2]Institute of Data Science, Maastricht University, Maastricht, The Netherlands
[3]Department of Radiation Oncology (MAASTRO), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre+, Dr Tanslaan 12, 6229ET, Maastricht, The Netherlands

ABSTRACT

In this paper, the authors describe an approach for sharing sensitive medical data with the consent of the data owner. The framework builds on the advantages of the Semantic Web technologies and makes it secure and robust for sharing sensitive information in a controlled environment. The framework uses a combination of Role-Based and Rule-Based Access Policies to provide security to a medical data repository as per the FAIR guidelines. A lightweight ontologywas developed, to collect consent from the users indicating which part of their data they want to share with another user having a particular role. Here, the authors have considered the scenario of sharing the medical data by the owner of data, say the patient, with relevant persons such as physicians, researchers, pharmacist, etc. To prove this concept, the authors developed a prototype and validated using the Sesame OpenRDF Workbench with 202,908 triples and a consent graph stating consents per patient.

KEYWORDS

Access Policies, Semantic Web, RDF/SPARQL, Role Based, Rule Based, FAIR, Consent

1. INTRODUCTION

Artificial intelligent based learning healthcare systems aim to use information technology and data infrastructures to rapidly apply scientific insights to clinical care and to power scientific discovery from clinical insights [1,2]. For this to work, large amounts of routine health care and scientific data need to be made FAIR - Findable, Accessible, Interoperable and Reusable [3] - for both humans and machines. The definition of FAIR principles isspecified by the GO FAIR initiative [30].

- Findable: Data that is easy to find and identify for computers and humans. It includes metadatato facilitate searching for specific datasets.
- Accessible: Easily retrievable and long-term data storage. It has well defined access protocols.
- Interoperable: Data that used in combination with other datasets without loss in meaning of terms and values.
- Reusable: Data that can be further processed for future research. Adequate information about the data procurement, consent agreements, processing methods (provenance) and license terms are necessary.

DOI: 10.5121/ijdms.2019.11302

However, access to clinical data is longstanding challenge, particularly given the administrative, political and ethical barriers [4].

Ethical challenges often center on the need to protect the privacy of patients. One ethically and legally accepted way of accessing and processing personal data, such as patient data, is to ask consent of data subject involved [5]. Such consent must be specific. In the health care context, this means that patients should be able to control access to specific data elements to specific persons or machines for specific uses. Consent also needs to be dynamic as new data elements become available all the time and patients may change their minds and have a right to be forgotten in some jurisdictions [5]. On top of patient specific control, access policies may also be informed by institutional or national guidelines, like the USA HIPAA law [6], which defines specific data elements to be removed when data is shared with parties not involved in the direct care of patients, such as scientists. The Automatable Discovery and Access Matrix is an example of an initiative to encode such guidelines for consumption by machines [7]. Given these requirements, consent systems to get access to clinical data require a finely grained, multi-level (patient, institutional, etc.) and dynamic combination of authentication (which human or machine agent is accessing the data) and authorization (what data is accessed by which agent and for what reason).

Semantic Web technologies can be used to make data FAIR [8] and have been used as first implementations of a rapid learning health care system [9]. The Semantic Web, and its associated standards such as the RDF universal data model [10,11, 28], the SPARQL query language [12,13,25], and the OWL ontology language [14,15], aims to extend the Web from only consisting of human readable documents to a Web of machine understandable data [16]. However, the vision of the Semantic Web is mostly centered on open, linked data [17] so authentication and especially authorization have received relative modest attention.

The aim of this research is to develop a specific consent-based authorization scheme for clinical data using Semantic Web technology. The authors approached this by reviewing state of the art existing methods and techniques and comparing these with the proposed method.

2. RELATED WORK

Previous efforts in this domain include the work by Finin et al. [18] who represented a role-based control model into OWL. Their research focuses on the description of role concepts and their relationships (such as hierarchies).However, these investigators did not apply the roles in subsequent access on the Semantic Web. The author Büyükkılıç [24] discussed about rule-based control model, however the limitations of his work in the conversion of XML schema as per the standards.

Rishi KanthSaripalle et al.[27], have addressed the security and privacy at the knowledge level. They have proposed a Role Based Access Control Model to provide permissions to a RDF knowledge source. However, this approach would require administration and maintenance of roles and permissions by a knowledge/database administrator.

Gabillon and Letouzey[19] proposes to create security views on an RDF graph, similar to views managed by relational database administrators. In this approach, a graph is first created (using a SPARQL CONSTRUCT query) in which the security or access policy is applied. Subsequently, this graph is offered to the authorized user. They use a query based enforcement framework where each user specifies a rule for his existing RDF graph, which defines who can view the graph and what part of the graph can be accessed.

The drawback of their approach is that the query framework is cumbersome. It needs to be re-constructed for a new access, requiring maintenance by the graph administrator. This approach is not efficient for the use case wherein the patient maintains their own consent on data distributed across institutions and provides consent at multiple levels, eg consent for access of complete graph or consent for a node. The authors have built on the work of Gabillon et al. while addressing their limitations.

Sacco and Passant [20] present a lightweight ontology, the Privacy Preference Ontology to restrict access to a particular RDF data. The users have to first specify an access space indicating to which part of the data graph and to whom the restrictions have to be applied. Once the access space is defined, the users can specify the fine-grained access policies to the data belonging to a particular access space. Again, the use case of patients maintaining their own consents, the limitations of such policies is that if there is modification in the dataset, almost all the rules have to undergo modification.

K. Mohan and M. Aramudhan[21] have defined access policies using ontology-based approach, where they consider Object and Data properties for providing a secure access to personal health data stored in the cloud. They have not addressed the workflow of who generates the rules and the approach lacks a role based approach which makes it easier to define the access policies for healthcare data.

In [22] Hannes Muhleisen et al. define PeLDS (Policy enabled Linked Data Server), where they partition the dataset into different named graphs and create temporary view on those graphs by defining rules. The rules are defined using SWRL (Semantic Web Rule Language). Each rule from the access policy is attributed with an additional consequence to add the rule identifier to a global list of matched rules. If such a rule matches due to sufficient access rights for the current user, it will be added to this list. The list of rules is evaluated, and for every triple matching the data classifications contained in the rules consequence predicate list is copied from the dataset to the result graph. Rules depend on the structure of the dataset where a view of the graph is obtained and in case the structure is changed, all the rules have to be modified accordingly. The framework can further be extended to other types of data such as DICOM data converted into RDF format [23, 29].

The proposed framework address the shortcomings outlined in the above paragraphs. To validate this framework,a prototype of the access policy framework was developed. Institutions and organizations can use this, to share the sensitive RDF data, to specific persons having specific roles. The framework builds on the advantages of the Semantic Web technologies and makes it secure and robust for sharing sensitive information in a controlled environment, as per the Accessible and Re-usable guidelines of FAIR data sharing principles.

3. Methods

3.1 Consent Ontology and Graph

A consent ontology was analyzed by Gabillon [19] with which the consents of the user can be collected and stored, see Figure1. The classes in the ontology are

- Informed Consent Ontology (Consent Ontology): This is the main class that defines the framework for collection and storage of consents.
- Role: This class specifies the role of the user requesting the data and the user from whom the data is requested.
- AccessPolicy: This class defines the Access Policy defined by the users who would like share their data in a restricted way.
- Action: Refers to the type of Action that a user can be perform while requesting the data. The Action may be SPARQL QUERY or a SPARQL UPDATE.
- Consent: This class defines the consents under an access policy defined by the user.

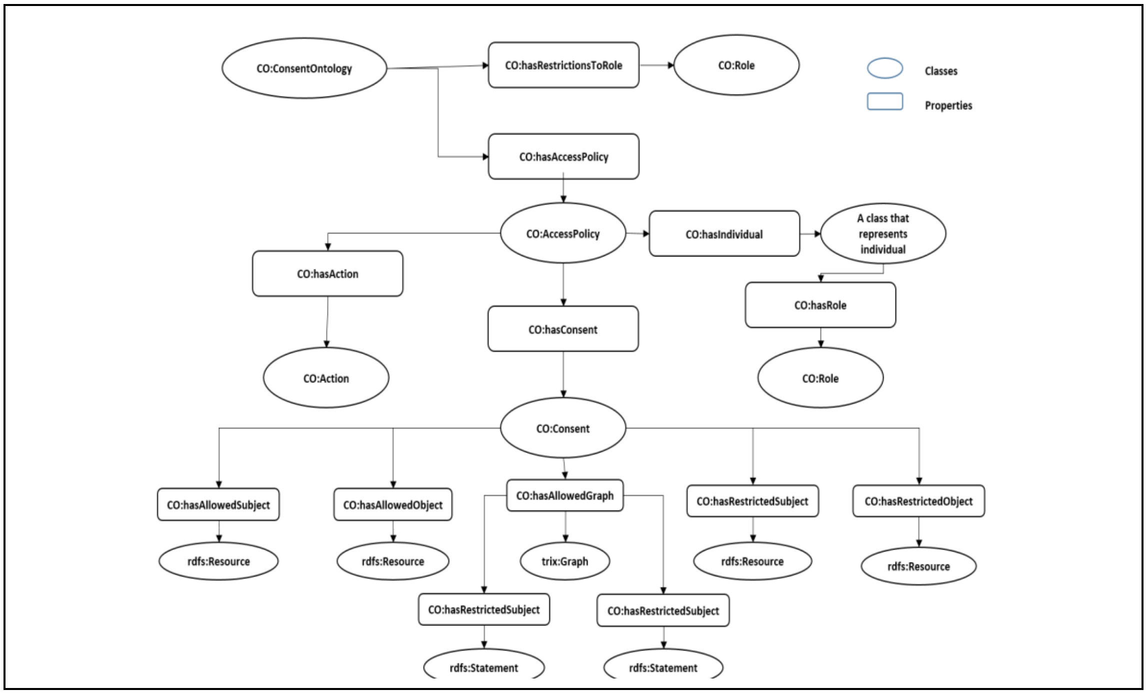


Figure 1 - The consent ontology used in the access policy framework

The consent ontology was used to define the consent graph for a specific patient. First, roles were defined as instances of the class ‘Role’ rather than as subclasses, as is recommended by Finin et al. [18]. Then an access policy is defined which includes the consents specified by a user for a given role. The specific consent specified in the example below which allows a researcher access to ICU data.

```
CO:ConsentOntology1 a CO:ConsentOntology ;
CO:hasRestrictionsToRoleCO:researcherRole.
CO:Access_Policy_1 a CO:AccessPolicy ;
CO:hasActionCO:Query ;
CO:hasConsent CO:Consent_1 ;
CO:hasIndividual<http://www.example.org/1> ;
CO:hasRoleCO:patientRole .
CO:Consent_1 a CO:Consent ;
CO:hasAllowedSubject<http://www.example.org/icu_data_1> .
CO:researcherRole a CO:Role ;
```

Figure 2 - Access Policy Framework

3.2 Access Policy Framework

An access policy framework was developed to consume the above described consents and allow users access to the data given their roles (Figure 2). When a user (e.g. a researcher) logs in and requests data from a given patient or patient cohort, a SPARQL CONSTRUCT query is created which creates a subgraph of the original data graph. This subgraph is then shared with the researcher for querying. More specifically, the CONSTRUCT query is a federated SPARQL query across the data graph and the consent graph with a filter for the role of the user. We first query the consent graph with graph identifier consentNamedGraph to obtain the consents. With the obtained consents from the consentNamedGraph, we query the patient graph with graph identifier patientNamedGraph to construct the graph with corresponding triples. Given below is the generic "patientToResearcherRule" using which queries are constructed by replacing patient_name, patientNamedGraph and consentNamedGraph with the URL of a specific patient, her patient graph identifier and the consent graph identifier (Figure 3).

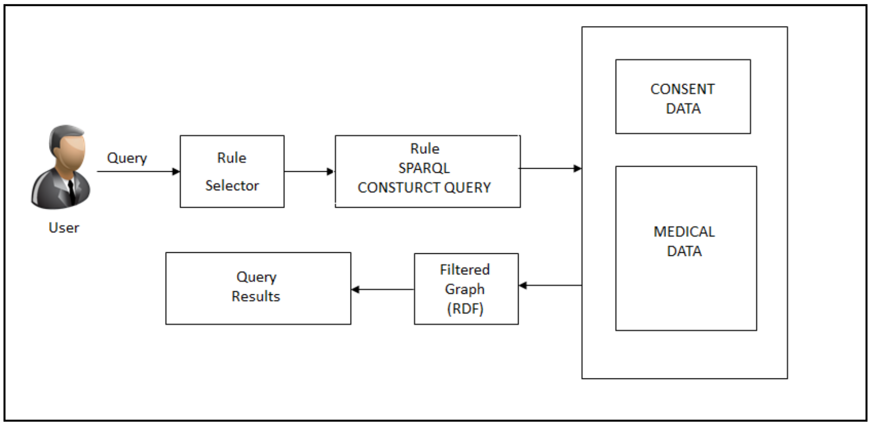


Figure 3 - Generic Rule

```
PREFIX BREASTCANCER: <http://breastcancer-data.org/>
PREFIX BREASTCANCERDOCUMENT: <http://breastcancer-data.org/document#>
PREFIX CO:
<http://www.semanticweb.org/310204290/ontologies/2016/3/ConsentOntology#>
PREFIX sedi: <http://semantic-dicom.org/dcm#>
PREFIX EMR: <http://emr-data.org/>
PREFIX DICOM: <http://dicom-data.org/>
PREFIX DICOMDOCUMENT: <http://dicom-data.org/document#>
PREFIX EMRDOCUMENT: <http://emr-data.org/document#>
PREFIX GENETIC: <http://genetic-data.org/>
PREFIX GENETICDOCUMENT: <http://genetic-data.org/document#>
PREFIX GRAPHIDENTIFIER: <http://www.namedGraph.org/>
PREFIX ICU: <http://icu-data.org/>
PREFIX ICUDOCUMENT: <http://icu-data.org/document#>
PREFIX OCKR: <http://www.example.org/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xml: <http://www.w3.org/XML/1998/namespace>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
//Comments: Based on Patient Name graph and Consent Name graph, constructing the
SPARQL Query to define access policy


CONSTRUCT{ ?s ?p ?o. }


FROM NAMED<patientNamedGraph>
FROM NAMED<consentNamedGraph>


WHERE {
  {
    GRAPH<patientNamedGraph>
  UNION
{ // subgraph is shared with the researcher for querying
   GRAPH<consentNamedGraph>
   {
     ?consentontologyCO:hasRestrictionsToRole ?role.
FILTER(?role = CO:researcherRole)
     ?role CO:hasAccess ?accessPolicy.
     ?accessPolicyCO:hasConsent ?consent.
     ?accessPolicyCO:hasIndividualOCKR:patient_name.
   }
   {
     GRAPH<consentNamedGraph>
     {
       ?consent CO:hasAllowedSubject ?s.
     }
     GRAPH<patientNamedGraph>
     {
       ?s ?p ?o.
     }
   }
   UNION
   {    //query the consent graph with graph identifier consentNamedGraph to obtain the
consents
     GRAPH<consentNamedGraph>
     {
       ?consent CO:hasAllowedSubject ?subject.
     }
     GRAPH<patientNamedGraph>
  UNION
  {
   GRAPH<consentNamedGraph>
   {
     ?consent CO:hasAllowedProperty ?p.
   }
        //query the patient graph with graph identifier patientNamedGraph
   GRAPH<patientNamedGraph>
   {
     ?s ?p ?o.
   }
  }
  UNION
  {     GRAPH<consentNamedGraph>
   {
     ?consent CO:hasAllowedProperty ?property.
   }
   GRAPH<patientNamedGraph>
   {
     ?s ?property ?o1.
VALUES ?p {rdf:type}
     ?s ?p ?o.
   }
  }
  UNION
  {
   GRAPH<consentNamedGraph>
   {
     ?consent CO:hasAllowedProperty ?property.
   }
```

```
        GRAPH<patientNamedGraph>
      {
        ?s1 ?property ?s.
VALUES ?p {rdf:type}
        ?s ?p ?o.
      }
    }// construct the graph with corresponding triples
    UNION
    {
      GRAPH<consentNamedGraph>
      {
        ?consent CO:hasAllowedGraph ?graph.
GRAPH ?graph {?s ?p ?o.}
      }
    }
  }
}
```

The given query efficiently retrieves the allowed graph or a combination of allowed subjects and/or allowed properties for a single patient. In order to retrieve allowed graphs for multiple patients from a triple-store, multiple queries need to be constructed from the patient To Researcher Rule and queried from the triple-store [25].

Before applying the access policies, the data of multiple patients is not stored in a single RDF graph. To enable retrieving data from multiple users, the complete data is partitioned into different datasets using named graphs. In this way, a graph identifier gets assigned to each triple, thus allowing easy retrieval of all triples belonging to a specific graph. Each dataset with a named graph has an owner who manages the data. In this case, each patient owns a RDF named graph.

4. EVALUATION

A number of real world use cases were used to evaluate the access policy framework. These use cases were

- A patient gives consent to the use of her birth data to a user with the role “Physician” but not to a user with the role “Researcher”.

- A user with the role “Researcher” request the cohort of patients who have consented to the use of a combination of subjects such as EMR data, breast cancer data, ICU data, genetic data or the entire graph - patientNamedGraph.
- A patient gives consent to a specific named user with the role “Physician” but not to another named user with the role “Physician”. (e.g. a patient may have a conflict with a certain physician)
- A patient gives consent but the data holder withholds consent (e.g. a patient may give consent to share a physician name with the outside world, which the hospital does not allow)

To evaluate if the access policy framework can fulfill these use cases a number of public datasets containing ICU, HIV and breast cancer data were used, (see Appendix and supplemental material). Using these we constructed 101 distinct patient graphs with 202,908 triples.

To specify the consents of all the patients, random subjects were selected and/or properties and/or and/or objects or the entire graph to be allowed by each patient and constructed the consent graph (815 triples, see supplemental material).

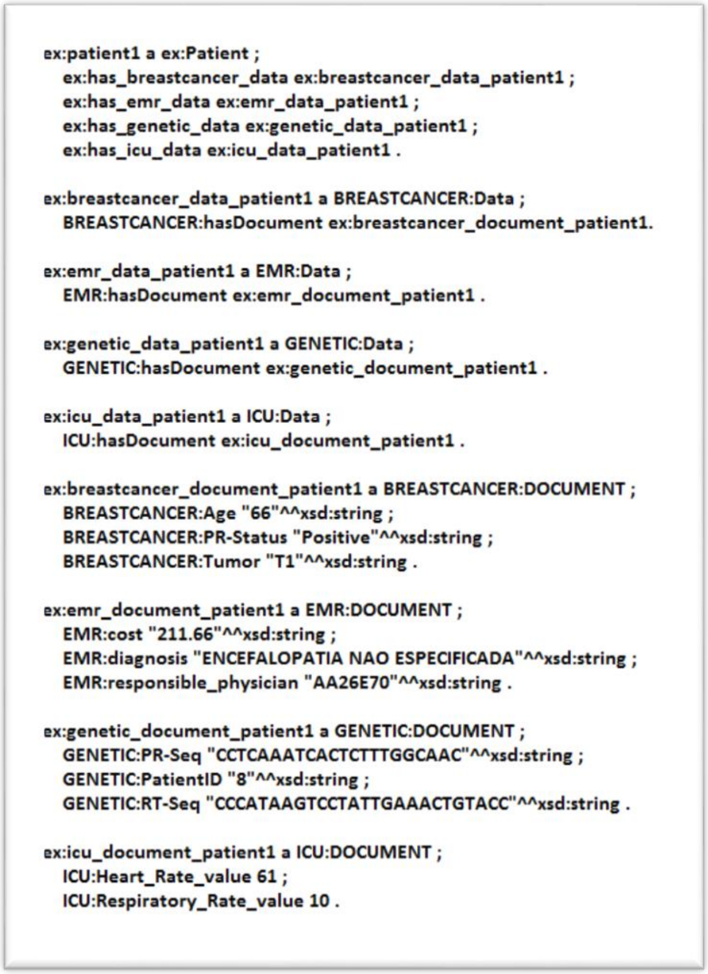
All data triples (n~200000) were loaded into Sesame OpenRDF-Workbench (version 4.1.2, Eclipse Foundation) on a Tomcat server (version 7.0, Apache Software Foundation). The evaluation was conducted on a computer with a processor (i5-4310-2GHz, Intel) and 8 GB of RAM. In each case, the SPARQL CONSTRUCT query consuming the consent ontology was executed first to obtain the filtered subgraph for each of the patients.

The procedure is as follows:

- The consent graph and patient graphs are loaded into the triple-store.
- Then a single query and retrieval is done for each patient graph leading to same number of queries as the number of patient graphs.

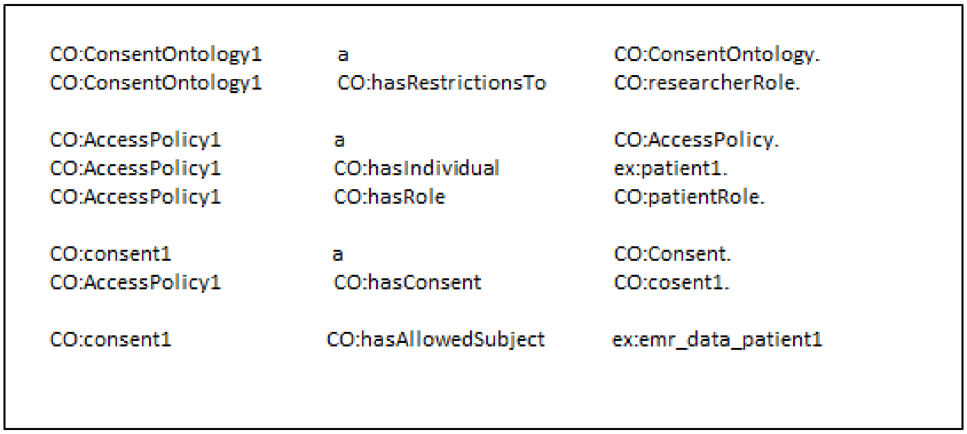
5. Results

For the purpose of demonstrating our Access Policy Framework, the authors have considered a RDF graph with a patient graph identifier as “ex:graph_patient1” and consent graph identifier as “ex:graph_consent” using synthetic medical data. Figure 4 and Figure5 shows a RDF document of a patient with ICU acquired data, EMR data, genetic information, breast cancer data and Consent Ontology.



```
ex:patient1 a ex:Patient ;
  ex:has_breastcancer_data ex:breastcancer_data_patient1 ;
  ex:has_emr_data ex:emr_data_patient1 ;
  ex:has_genetic_data ex:genetic_data_patient1 ;
  ex:has_icu_data ex:icu_data_patient1 .

ex:breastcancer_data_patient1 a BREASTCANCER:Data ;
  BREASTCANCER:hasDocument ex:breastcancer_document_patient1.

ex:emr_data_patient1 a EMR:Data ;
  EMR:hasDocument ex:emr_document_patient1 .

ex:genetic_data_patient1 a GENETIC:Data ;
  GENETIC:hasDocument ex:genetic_document_patient1 .

ex:icu_data_patient1 a ICU:Data ;
  ICU:hasDocument ex:icu_document_patient1 .

ex:breastcancer_document_patient1 a BREASTCANCER:DOCUMENT ;
  BREASTCANCER:Age "66"^^xsd:string ;
  BREASTCANCER:PR-Status "Positive"^^xsd:string ;
  BREASTCANCER:Tumor "T1"^^xsd:string .

ex:emr_document_patient1 a EMR:DOCUMENT ;
  EMR:cost "211.66"^^xsd:string ;
  EMR:diagnosis "ENCEFALOPATIA NAO ESPECIFICADA"^^xsd:string ;
  EMR:responsible_physician "AA26E70"^^xsd:string .

ex:genetic_document_patient1 a GENETIC:DOCUMENT ;
  GENETIC:PR-Seq "CCTCAAATCACTCTTTGGCAAC"^^xsd:string ;
  GENETIC:PatientID "8"^^xsd:string ;
  GENETIC:RT-Seq "CCCATAAGTCCTATTGAAACTGTACC"^^xsd:string .

ex:icu_document_patient1 a ICU:DOCUMENT ;
  ICU:Heart_Rate_value 61 ;
  ICU:Respiratory_Rate_value 10 .
```

Figure 4- RDF Document:ICU, EMR, genetic and breast cancer data



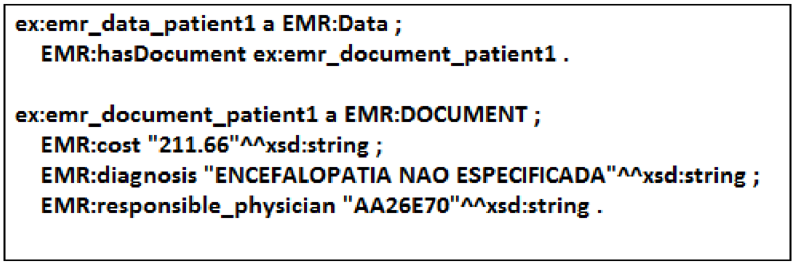
```
CO:ConsentOntology1     a                      CO:ConsentOntology.
CO:ConsentOntology1     CO:hasRestrictionsTo   CO:researcherRole.

CO:AccessPolicy1        a                      CO:AccessPolicy.
CO:AccessPolicy1        CO:hasIndividual       ex:patient1.
CO:AccessPolicy1        CO:hasRole             CO:patientRole.

CO:consent1             a                      CO:Consent.
CO:AccessPolicy1        CO:hasConsent          CO:cosent1.

CO:consent1             CO:hasAllowedSubject   ex:emr_data_patient1
```

Figure 5 - Consent Ontology

The authors consider a scenario in which the medical data of a patient in a hospital has to be shared with a researcher. The authors evaluated the proposed access policies for three different cases in which the consents of the patient given to a researcher are different. For each of the cases, the rules applied were the same as described in the METHODS section.

Case 1: Patient shares only the EMR data.

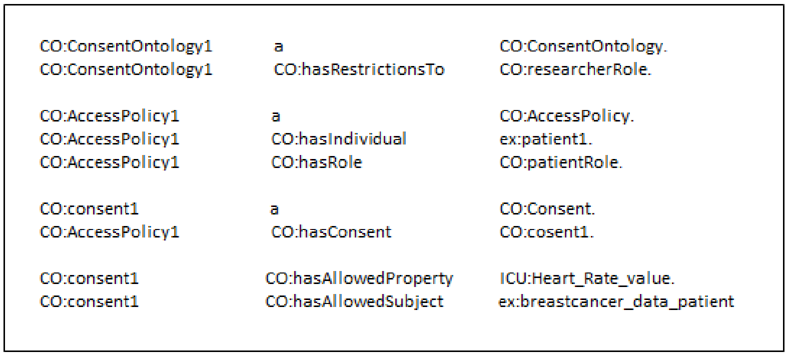
The consent graph for this case is shown in Figure 6, where the patient has allowed the subject ex:emr_data_patient1, which corresponds to the EMR data of the patient.



```
ex:emr_data_patient1 a EMR:Data ;
   EMR:hasDocument ex:emr_document_patient1 .

ex:emr_document_patient1 a EMR:DOCUMENT ;
   EMR:cost "211.66"^^xsd:string ;
   EMR:diagnosis "ENCEFALOPATIA NAO ESPECIFICADA"^^xsd:string ;
   EMR:responsible_physician "AA26E70"^^xsd:string .
```

Figure 6- Consent Graph EMR Data

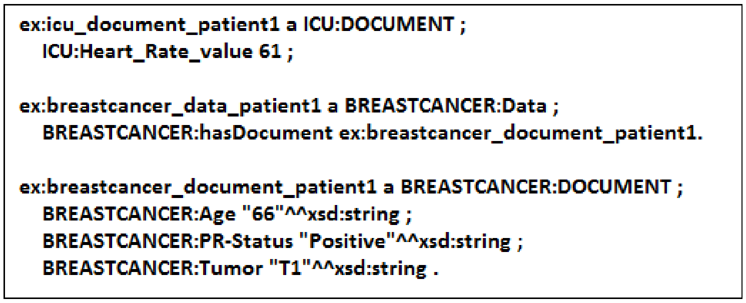
Figure 7 shows the filtered graph obtained after the rule is applied on the consents and the patient data. The filtered graph is a sub graph of the patient data graph starting from the subject specified in the consents.



```
CO:ConsentOntology1      a                         CO:ConsentOntology.
CO:ConsentOntology1      CO:hasRestrictionsTo      CO:researcherRole.

CO:AccessPolicy1         a                         CO:AccessPolicy.
CO:AccessPolicy1         CO:hasIndividual          ex:patient1.
CO:AccessPolicy1         CO:hasRole                CO:patientRole.

CO:consent1              a                         CO:Consent.
CO:AccessPolicy1         CO:hasConsent             CO:cosent1.

CO:consent1              CO:hasAllowedProperty     ICU:Heart_Rate_value.
CO:consent1              CO:hasAllowedSubject      ex:breastcancer_data_patient
```

Figure 7 - Filtered Graph post application of consents

Case 2: Patients shares only heart rate in the ICU data and the Breast cancer data shown in figure 8.

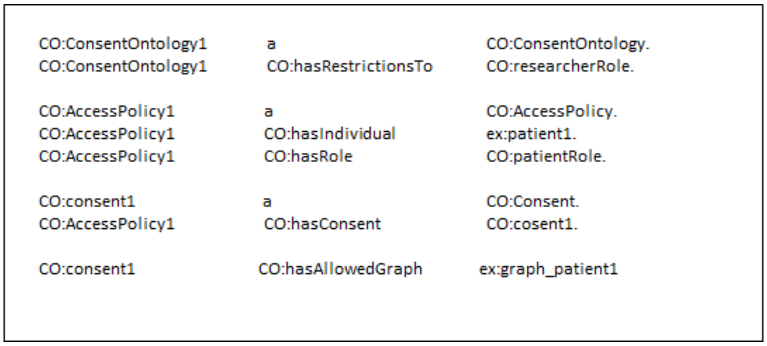


```
ex:icu_document_patient1 a ICU:DOCUMENT ;
   ICU:Heart_Rate_value 61 ;

ex:breastcancer_data_patient1 a BREASTCANCER:Data ;
   BREASTCANCER:hasDocument ex:breastcancer_document_patient1.

ex:breastcancer_document_patient1 a BREASTCANCER:DOCUMENT ;
   BREASTCANCER:Age "66"^^xsd:string ;
   BREASTCANCER:PR-Status "Positive"^^xsd:string ;
   BREASTCANCER:Tumor "T1"^^xsd:string .
```

Figure 8 - Heart rate in ICU Data and Breast Cancer Data

Case 3: Patient shares the complete data (Full graph)

Figure 9 shows the consent given by the patient to share the complete graph. There the complete graph with the specific graph identifier will be constructed as a result and shared with the researcher.



```
CO:ConsentOntology1     a                    CO:ConsentOntology.
CO:ConsentOntology1     CO:hasRestrictionsTo CO:researcherRole.

CO:AccessPolicy1        a                    CO:AccessPolicy.
CO:AccessPolicy1        CO:hasIndividual     ex:patient1.
CO:AccessPolicy1        CO:hasRole           CO:patientRole.

CO:consent1             a                    CO:Consent.
CO:AccessPolicy1        CO:hasConsent        CO:cosent1.

CO:consent1            CO:hasAllowedGraph   ex:graph_patient1
```

Figure 9- Patient Consent to share full graph

The total query execution time for all the patients was 111.64 seconds with 70,190 triples constructed as a part of subgraph or the filtered graph. The total query execution time refers to get all the filtered graphs for the patient in the triple store. The measurements shows a significant improvement over standard SQL based query and retrieve methods.

6. DISCUSSION

The framework uses a combination of Role Based and Rule Based Access Policies to provide security to a medical data repository. The prototype is validated using Sesame OpenRDF Workbench with 202,908 triples and a consent graph stating consents per patient. The main advantage of this Access Policy being, there is no requirement for each user to specify the rules. The user will only have to provide the consents. The Central Authority or an Administrator who is officially responsible for the medical database can specify the rules. The rules are specified according to the structure of the data, irrespective of the consents given by the users.

Various authors have presented their work on access policies for the sensitive RDF data. Finin et al. [18] integrated RBAC (Role Based Access Control) model into OWL. They used OWL ontologies to represent RBAC model which specifies the access control policies. Gabillon and Letouzey[19] emphasize on providing access control policies over named graphs and views which generates a subgraph. They use query based enforcement framework where each user specifies a rule for his existing RDF graph, which defines who can view the graph and what part of the graph can be accessed. Sacco and Passant [20] present a light weight ontology, Privacy Preference Ontology(PPO) to restrict access to a particular RDF data. The users have to first specify an access space indicating to which part of the data graph and to whom the restrictions have to be applied. Once the access space is defined, the users can specify the fine-grained access policies to the data belonging to a particular access space. The limitations of such policies is that if there is modification in the dataset, almost all the rules have to undergo modification.

However, in each of the approaches, users have to define their own rule which becomes cumbersome and difficult to manage. Also, in organizations like hospital, if each patient defines their own rules, it leads to duplication of the rules since the structure of the dataset remain the same. Instead, a single rule can be defined for all the patients with only customizations in consents.

Artificial intelligence, machine-learning algorithms require training on large sets of curated data. In a typical scenario, to train an algorithm, the researcher would obtain consent from the data owner; protect the privacy, via de-identification [31] as the initial process before using the data for further analysis. When the same researcher requires the same data for a different purpose, the researcher maybe required to take another consent from the data owner. Currently, the method and techniques for seeking consent and maintaining a record is semi-automated and in many cases manual. The data owner may not be fully cognizant of the consent process and its real world deployment. Similarly, the researcher finds it challenging to obtain data, which has the necessary consent approvals and adheres to the country specific privacy and security laws. The re-usability of data per the FAIR guidelines, enabling data usage for future research and further computational processing with satisfactory licenses and provenance, is an area requiring improvement with process, tools and techniques.

This framework is a step in the direction of empowering the data owner to manage the rights of data usage. Extending this concept further, the framework can include audit trails, history, duration for data usage and purpose of usage.

7. CONCLUSION

The authors built a lightweight ontology to collect the consents from the users indicating which data entities they want to share with another user with a specific role. To address, concerns of security and privacy of medical data, the authors considered the scenario of sharing the medical data of a hospital amongst different stakeholders such as a patient, physician, and a researcher. The authors have followed a hybrid approach with a combination of Role-based and Rule-based Access to provide security to a medical RDF data. A central authority, knowledgeable in semantic web and the database construct the rules, and the patients provide the consents. The advantages of the authorization framework is the ease by which a user can change access rights, by modifying only the consents. Similarly, when there is an update to the dataset, only the rules specified by central authority requires modification.

8. ACKNOWLEDGEMENTS

The authors thank the professors at Maastricht University and Dr. Shyam Vasudev Rao, Dean of Maastricht University, India for their support and valuable guidance. They acknowledge the contributions of colleagues at Philips Research, Bangalore, specially Ravindra Patil and Kiran Kumar for their valuable review comments.

REFERENCES

[1] Maddox TM, Albert NM, Borden WB, Curtis LH, Ferguson TB, Kao DP, et al. The Learning Healthcare System and Cardiovascular Care: A Scientific Statement From the American Heart Association. Circulation. 2017;135:e826–57.

[2] Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CML, et al. ‘Rapid Learning health care in oncology’ – An approach towards decision support systems enabling customised radiotherapy’. Radiother. Oncol. 2013;109:159–64

[3] Wilkinson MD, Dumontier M, AalbersbergIjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data. 2016;3:160018.

[4] Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol NJ, Amir E, et al. Delivering affordable cancer care in high-income countries. Lancet Oncol. 2011;12:933–80.

[5] Regulation GDP. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46. Off. J. Eur. Union OJ. 2016;59:1–88.

[6] Standards for privacy of individually identifiable health information. Office of the Assistant Secretary for Planning and Evaluation, DHHS. Final rule. Fed. Regist. 2000;65:82462–829.

[7] Automatable Discovery and Access Matrix (“ADA-M”)v1.0-Guidance Document[Internet].Global Alliance for Genomics & Health (GA4GH) and International Rare Disease Research Consortium (IRDiRC); http://genomicsandhealth.org/files/public/ADAM_GuidanceDocument_15Dec2016_Final.pdf

[8] Wilkinson MD, Verborgh R, Bonino da Silva Santos LO, Clark T, Swertz MA, Kelpin FDL, et al. Interoperability and FAIRness through a novel combination of Web technologies. PeerJComput. Sci. 2017;3:e110.

[9] Jochems A, Deist TM, Soest J van, Eble M, Bulens P, Coucke P, et al. Distributed learning: Developing a predictive model based on data from multiple hospitals without data leaving the hospital – A real life proof of concept. Radiother. Oncol. 2016;121:459–67.

[10] RDF Primer [Internet]. [cited 2017 Jun 8]. Available from: https://www.w3.org/TR/2004/REC-rdf-primer-20040210/

[11]Decker S, Mitra P, Melnik S. Framework for the semantic Web: an RDF tutorial. IEEE Internet Comput. 2000;4:68–73.

[12] SPARQL Query Language for RDF [Internet]. [cited 2017 Jun 8]. Available from: https://www.w3.org/TR/rdf-sparql-query/

[13] Arenas M, Pérez J. Querying semantic web data with SPARQL. Proc. Thirtieth ACM SIGMOD-SIGACT-SIGART Symp. Princ. Database Syst. [Internet]. ACM; 2011 [cited 2017 Jun 8]. p. 305–316. Available from: http://dl.acm.org/citation.cfm?id=1989312

[14] OWL 2 Web Ontology Language Primer (Second Edition) [Internet]. [cited 2017 Jun 9]. Available from: https://www.w3.org/TR/owl2-primer/

[15]Jonquet C, Shah N, Youn C, Callendar C, Storey M-A, Musen M. NCBO annotator: semantic annotation of biomedical data. Int. Semantic Web Conf. Poster Demo Sess. [Internet]. 2009 [cited 2017 Jun 8]. Available from: http://www.lirmm.fr/~jonquet/publications/documents/Demo-ISWC09-Jonquet.pdf

[16] Berners-Lee T, Hendler J. Publishing on the semantic web. Nature. 2001;410:1023–4.

[17] Bizer C, Lehmann J, Kobilarov G, Auer S, Becker C, Cyganiak R, et al. DBpedia-A crystallization point for the Web of Data. Web Semant. Sci. Serv. Agents World Wide Web. 2009;7:154–165

[18] Finin T, Joshi A, Kagal L, Niu J, Sandhu R, Winsborough W, et al. R OWL BAC: representing role based access control in OWL. Proc. 13th ACM Symp. Access Control Models Tec Symp. Access

Control Models Technol. [Internet]. ACM; 2008 [cited 2017 Jun 8]. p. 73–82. Available from: http://dl.acm.org/citation.cfm?id=1377849hnol. [Internet]. ACM; 2008 [cited 2017 Jun 8]. p. 73–82. Available from: http://dl.acm.org/citation.cfm?id=1377849

[19] Gabillon A, Letouzey L. A View Based Access Control Model for SPARQL. IEEE; 2010 [cited 2017 Jun 8]. p. 105–12. Available from: http://ieeexplore.ieee.org/document/5636084/

[20] Sacco O, Passant A. A Privacy Preference Ontology (PPO) for Linked Data. LDOW [Internet]. 2011 [cited 2017 Jun 8]. Available from: http://www.academia.edu/download/6230668/ldow2011-paper01.pdf

[21] Mohan K, Aramudhan M. Ontology based access control model for healthcare system in cloud computing. Indian J. Sci. Technol. 2015;8:218–222.

[22] Muhleisen H, Kost M, Freytag J-C. SWRL-based access policies for linked data. Procs SPOT [Internet]. 2010 [cited 2017 Jun 8];80. Available from: http://ceur-ws.org/Vol-576/paper1.pdf

[23] Mahadevaiah G, Soest JV, Dekker A, Udupa N, Rao SV, Kumar YK, et al. Semantic Representation of Radiotherapy data for effective data mining. Proc. Fifth Int. Conf. Adv. Appl. Sci. Environ. Eng. - ASEE 2016 [Internet]. Kuala Lumpur, Malaysia: Institute of Research Engineers and Doctors, USA; [cited 2017 Jun 8]. p. 12–5. Available from: http://www.seekdl.org/nm.php?id=7421

[24] BÜYÜKKILIÇ, T .RULE-BASED SEMANTIC STANDARDS: A CONCEPTUAL FRAMEWORK FOR RULE-BASED SEMANTIC IS STANDARDS DEVELOPMENT, 2011.

[25] http://ai.ia.agh.edu.pl/wiki/_media/pl:dydaktyka:semantic_web:sparql.pdf

[26] Hira Asghar, Zahid Anwar, Khalid Latif: A deliberately insecure RDF-based Semantic Web application framework for teaching SPARQL/SPARUL injection attacks and defense mechanisms. Computers & Security 58: 63-82 (2016)

[27] Rishi KanthSaripalle, Alberto De la Rosa Algarin, Timoteus B. Ziminski: Towards knowledge level privacy and security using RDF/RDFS and RBAC. ICSC 2015: 264-267

[28] CsillaFarkas, VaibhavGowadia, Amit Jain, D. Roy: From XML to RDF: Syntax, Semantics, Security, and Integrity (Invited Paper). IICIS 2004: 41-55

[29] VaibhavGowadia, CsillaFarkas: RDF metadata for XML access control. XML Security 2003: 39-48

[30] Go FAIR Initiative :https://www.go-fair.org/fair-principles/

[31] ÖzlemUzuner, Yuan Luo, Peter Szolovits; “Evaluating the State-of-the-Art in Automatic De-identification.” Journal of the American Medical Informatics Association, Volume 14, Issue 5, 1 September 2007, Pages 550–563

Author

Geetha Mahadevaiah,Senior Director at Philips, Research Department, PIC, Bangalore.30+ years of experience in software engineering and management.*Areas of interest* : Clinical decision support systems, Semantic Web, healthcare applications Bachelor of Engineering in Computer Science and Technology Bangalore University Master of Business Administration, Bangalore University