A Policy Framework for Subject-Driven Data Sharing

Mohammad Jabed Morshed Chowdhury, Alan Colman, Jun Han
School of Software and Electrical Engineering
Swinburne University of Technology, Australia
{mjchowdhury,acolman,jhan}@swin.edu.au

Muhammad Ashad Kabir
School of Computing and Mathematics
Charles Sturt University, NSW, Australia
akabir@csu.edu.au

Abstract

Organizations (e.g., hospitals, university etc.) are custodians of data on their clients and use this information to improve their service. Personal data of an individual therefore ends up hosted under the administration of different data custodians. Individuals (data subjects) may want to share their data with others for various reasons. However, existing data sharing mechanisms provided by the data custodians do not provide individuals enough flexibility to share their data, especially in a cross-domain (data custodian) environment. In this paper, we propose a data sharing policy language and related framework for a data subject to capture their fine-grained data sharing requirements. This proposed language allows the data subject to define data sharing policies that consider context conditions, privacy obligations and re-sharing restrictions. Furthermore, we have implemented a prototype to demonstrate how data subjects can define their data sharing policies and how the policies can be used and enforced at runtime.

1. Introduction

Nowadays most of the traditional organizations (e.g., hospital, university, bank etc.) use web enabled systems to provide their services and host their clients' data. Individuals also use Web 2.0 platforms such as online social networks, blogs, etc. to share their personal data in the form of thoughts, photos, videos and ideas. Furthermore, pervasive and mobile computing applications are dramatically increasing the amount of personal data under the control of different service providers (data custodians). Different data custodians have different goals and policies related to the individual data subject's access to, and control over, the personal data about them. In particular, data custodians often lack the proper mechanisms to enable an individual to share their own data with other individuals or organizations – they become the "silos" of information for any individual.

Such personal information is typically subject to agreements or legislation related to privacy. Governments around the world are providing legal rights on the personal data hosted at different data custodians to their citizen. For example, Finland's MyData project [1] is a human centered approach in personal data management that combines industry needs with digital human rights. The core idea is to let individuals be in control of their own data. The General Data Protection Regulation (GDPR) adopted by the European Parliament aims to strengthen and unify data protection for individuals within the European Union (EU) [2]. This regulation mandates the data custodian to take individual's consent for data collected and purposes data used, and obliges the data custodian to use "Privacy by Design" principles [3]. "Privacy by Design" principles emphasize on User-Centric privacy when dealing with individual's data. Government of UK and Australia also formulated laws which are more directly related to data sharing, such as "Data sharing code of practice" [4] and "Guidelines for sharing personal information" [5].

Recently, online social networks (e.g., Facebook, Twitter, and LinkedIn) have revolutionized the individual's attitude towards their personal data sharing. People share their personal information, photos, and videos with friends, family members and colleagues in online social networks. Individuals may also want to provide other individuals secure access to their personal data hosted in different data custodian systems. This might be for personal, social, medical, financial or other reasons. They may also wish to share their personal data with different service providers to get better and/or personalized services. The data subject may also wish to provide data consumer limited rights to re-share their data. In this paper we use the term "data sharing" to mean providing selective and conditional access (not providing a copy of the original data) to

URI: http://hdl.handle.net/10125/50483
ISBN: 978-0-9981331-1-9
(CC BY-NC-ND 4.0)

a data resource held by a data custodian about an individual to individuals outside of the data custodian system(which is hosting the data). Our assumption in this work is that the data consumer is granted access (to the data resource hosted at data custodian) to ensure the authenticity of the data resource (as data is being accessed from the source (e.g., from the university server)). We also assume the data subject trusts the data consumer not to violate agreed privacy provisions (e.g. unauthorised sharing or copying). Mechanisms for preventing violation of trust typically rely on the consumer giving a legally enforceable undertaking via a conditions of use agreement prior to accessing the data. Enforcing compliance with such agreements can be supported by access audit-trails, watermarks, legal remedies, etc., however discussion of mitigating such breach of trust is outside the scope of this paper.

Although there is a clear need for individuals to share their personal data and there is legislative impetus to support this, current data access policy languages and mechanisms do not support individuals' fine grained control over their data sharing. Data custodians lack the mechanisms to empower data subjects to define their own personal data sharing policies that capture their privacy preferences in various contexts. Access-control can be expressed in languages like XACML [6] and enforced by the enterprise data custodian. Likewise, individual's privacy requirements and re-sharing preferences need to be expressed and enforced by either technical measures or legal sanctions. Current access control policy languages does not have enough constructs to capture all contextual, privacy preserving and re-sharing requirements (details in section 2) of individuals. Thus, we need a data sharing policy language that can capture individual's privacy preference and re-sharing preference along with the contextual access control requirements over her data sharing.

The essence of our approach is to capture an individual's data sharing requirements related to contextual conditions, privacy obligations, re-sharing rights and constraints. We analyze a data sharing scenario (presented in section 2). We follow resource-centric approach in this work where each sharing resource will be associated with its data sharing policy (ies) defined primarily by the data subject (subject to any administrative policies defined by the data custodian). The contributions of this paper are to differentiate the key roles when sharing access to data, and to specify a data sharing policy language which enables the data subject to specify sharing, privacy and contextual constraints over access personal data.

In section 2, we present a data-sharing scenario and then analyze the scenario to elicit the types of requirement an individual may have when sharing his/her data. Section 3 presents our data sharing policy language. Section 4 presents three case studies to validate our proposed data sharing policy. Section 5 presents an access-control domain model which supports data sharing policies and an implementation of the framework. Section 6 discusses the related work. We conclude the paper with a discussion of future work in Section 7.

2. Scenario and Analysis

In this section, we present a scenario that illustrates a case of an individual controlling the privacy and sharing their data with third party. We define some key terms and analyze the scenario with respect to data sharing.

2.1. Scenario

Alice was a student at ABC University, who has recently graduated. Students can securely access their own academic transcripts in a number of data formats. Transcripts are also available in detailed (all results) and summary (GPA or CGPA). She has applied for a job position at Company XYZ. Her prospective employer Mr. Smith has liked her application and wanted to see her original transcript. In the past, he has been presented with real-looking but fake copies from other applicants and he wants to ensure the transcript is genuine. He therefore wants to access the transcript directly from the University itself. Alice is happy to share access to her academic transcript with Mr. Smith, however she has some privacy concerns over this sharing: Alice wants:

1. Mr. Smith from Company XYZ to be able to access to her academic record hosted at ABC University.
2. Mr. Smith to have access only from 01/06/2017 to 10/06/2017.
3. Access to her record can only be gained from the XYZ corporate office.
4. An undertaking that her academic record can only be used for purposes related to her job application.
5. To allow Mr. Smith to be able to delegate "sharing rights" so that he can share with others in XYZ for the purposes of the job application.
6. Mr. Smith should not be able to share it outside Company XYZ employee.

7. If Mr. Smith shares her transcript then the new data consumer should only see her Cumulative Grade point Average (CGPA) rather than detailed transcript.
8. Mr. Smith can only share it with 3 persons and the re-sharing data consumer cannot share it further.
9. When someone (Mr. Smith or anyone else) accesses her academic record the data custodian keeps a record of the share event and sends her an email notification.

The presented data-sharing scenario highlights different types of requirement Alice has over her academic record sharing. The general characteristics of this sharing is that Alice wants enough flexibility to define all her data sharing requirements and wants to share her data with third parties who are not part of the same data custodian which is hosting the data resource. Though Alice has different types of requirements, we can categorize them into the following three broad categories: i. Context Conditions, ii. Privacy Obligations, and iii. Re-Sharing Conditions. In the next section we will provide terminologies related to these categories.

2.2. Terminologies

Figure 1 shows the data sharing scenario using our proposed terminologies. The *data subject* (Alice) has the right to access a *data resource* (her academic record) hosted at a *data custodian* (ABC University). She defines a *data sharing policy* to share the resource with a *data consumer* (Mr. Smith). When the *data consumer* wants to access that shared resource he must have to satisfy all the conditions (e.g., contextual, privacy obligations, and re-sharing conditions) before getting access to that particular resource. These terms used in our proposed *data sharing policy language* are defined as follows: *Data Subject:* is an individual whose

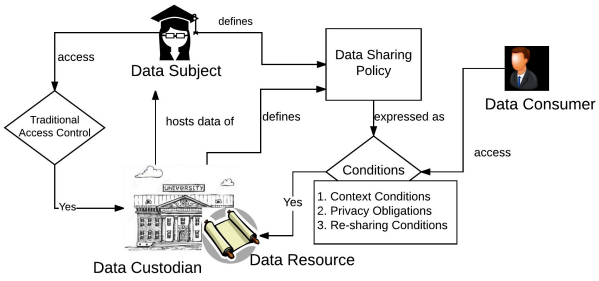

Figure 1. Data Sharing Scenario.

personal information is hosted in the data custodian system. *Data subjects* write *data sharing policy* to share their personal data. They can also delegate their "sharing rights" to the *data consumer*.

Data Resource: any information about the data subject that might be outside the public domain or subject to privacy restrictions.

Data Consumer: is an individual who is granted access to data subject's data resource. The data subject defines data sharing policy to share data resource with the data consumer. If the data consumers is granted "sharing rights", then they can re-share that resource with other data consumer.

Data Custodian: hosts the data resource about the data subject. They provide the mechanism to the data subject to define the data sharing policy. Data custodian is also responsible to enforce the policy defined by the data subject. Data custodian defines organizational policy related to the data in accord with any regulatory or legislative constraints.

Context Conditions: are defined in the policy to ensure context-aware access to the data resource. Context is defined by the attributes of the data subject, data consumer or environmental state. Context could be related to location, time, social relationship, physical condition or any other conditions needed by the data subject to provide selective conditional sharing.

Privacy Obligations: The following terminologies are related to privacy obligation of the data subject.

Purpose: captures any restriction the data subject wants to put on the data consumer related to reasons for accessing their data, or restrictions on what use that is made of the data. Purpose cannot typically be enforced by the custodian's system, but is a (potential legally enforceable) obligation agreed to by the data consumer.

Data Representation: Individuals may wish to share different types of data representation to different types of data consumer. There can be different types of data representation requirements, such as (i) granularity of data (e.g. Alice's full record or GPA only), ii) anonymization, and iii) encryption of data. Responsibility for providing data at a range granularity, anonymization and encryption lie with the data custodian. The policy language we propose enables the data subject to state their requirements in terms of the various representations provided by the custodian.

Accounting and Notification: The data sharing code of practice [4] requires that the data custodian keep a record of the access. There may also be a requirement the data subject be notified when their data is accessed.

Re-sharing Conditions: The following terminologies are related to privacy obligation of the data subject. The following terminologies are related to re-sharing requirements of the data subject.

Sharing Rights Delegation: When a data subject shares a data resource with the data consumer s/he may want to delegate his/her "sharing rights" to the data consumer so that s/he (data consumer) can share it further with other data consumer.

Re-sharing Conditions: When data subject delegates "sharing rights" to the data consumer, s/he may wish to impose some constraints on re-sharing of data by the data consumer. Data re-sharing constrains is expressed as a data sharing policy that additionally captures the following two requirements

i. Cardinality of Sharing: Maximum number of times (data consumers) a resource can be concurrently shared.

ii. Recurring Sharing: Maximum number of steps (data consumer 1 $\rightarrow$ data consumer 2 $\rightarrow$ data consumer 3) that a data resource can be re-shared.

2.3. Requirements Analysis

From an analysis of the scenario, we define the following requirements of the proposed policy language.

Requirement 1: The policy language should be able to specify sharing policy in a resource-centric manner. The *data subject* should be able to share their data resource with anyone outside of the data custodian's system. The resource should be addressable in a way that can be interpreted universally. In addition, every shared resource should have its own data sharing policy/ies associated with it. When access is requested to any particular shared resource only the related/attached policy(ies) will be evaluated.

Requirement 2: The language should allow data subject to define contextual conditions for access. We have seen in the scenario that Alice wants to share her academic record but she puts some conditions related to the access control such as time and network context.

Requirement 3: The language should allow to define privacy requirements of data subject. Data sharing does not involve only access control. As personal data is being shared, individuals also may have privacy concerns related to their data sharing. The policy languages should be flexible enough to capture privacy preferences of data subject. For example, in the scenario Alice wants to share her academic record only for job applications purposes. The language should also allow us to define various representations (How should it be shared?[4]) of the resource related to data granularity, anonymisation or encryption and other privacy requirements.

Requirement 4: The language should allow data subject to define re-sharing rights and constraints. Data subjects may wish to delegate their data "sharing rights" to the data consumer to achieve a goal. For example, in the scenario Alice delegates her "sharing rights" to Mr. Smith so that he (Mr. Smith) can share it with his HR department for the paper work related to her (Alice) job application. In addition, sometime data subjects may wish to put some constraints on the re-sharing of their data because of security or privacy issues. For example, in the scenario, Alice wants to restrict Mr. Smith's "sharing rights" by putting a constraint like "can only be re-shared among the company (xyz.com) employees."

3. Policy Language

Existing access control (e.g., XACML[6]) and privacy policy languages (e.g., P3P [7], EPAL[8]) do not support all the requirements identified in the previous section. We have presented a brief description of the existing policy languages and their limitations in section 2. Our proposed data sharing policy language addresses these limitations in enabling data subjects to define their data sharing requirements.

3.1. Meta model of Data Sharing Policy Language

Figure 2 shows the meta model of our proposed data sharing policy language. In this section, we describe different concepts and their relationship in the proposed data sharing policy language. The proposed approach is resource-centric and the policies are attached to the specific Data Resource. Data Sharing Policy associated with any data resource will be evaluated when the access decision need to be made.

Data Sharing Policy is used to define data consumer, who can gain access to the data resource based on the *context condition*, *privacy obligation*, and *re-sharing condition* by the data subject. *Data Sharing Policy* has an associated class *Decision* that defines the effect of the policy as *permit* or *deny*.

ContextCondition is used to capture the contextual conditions related to the access of the data resource. ContextCondition is defined using the attributes of the

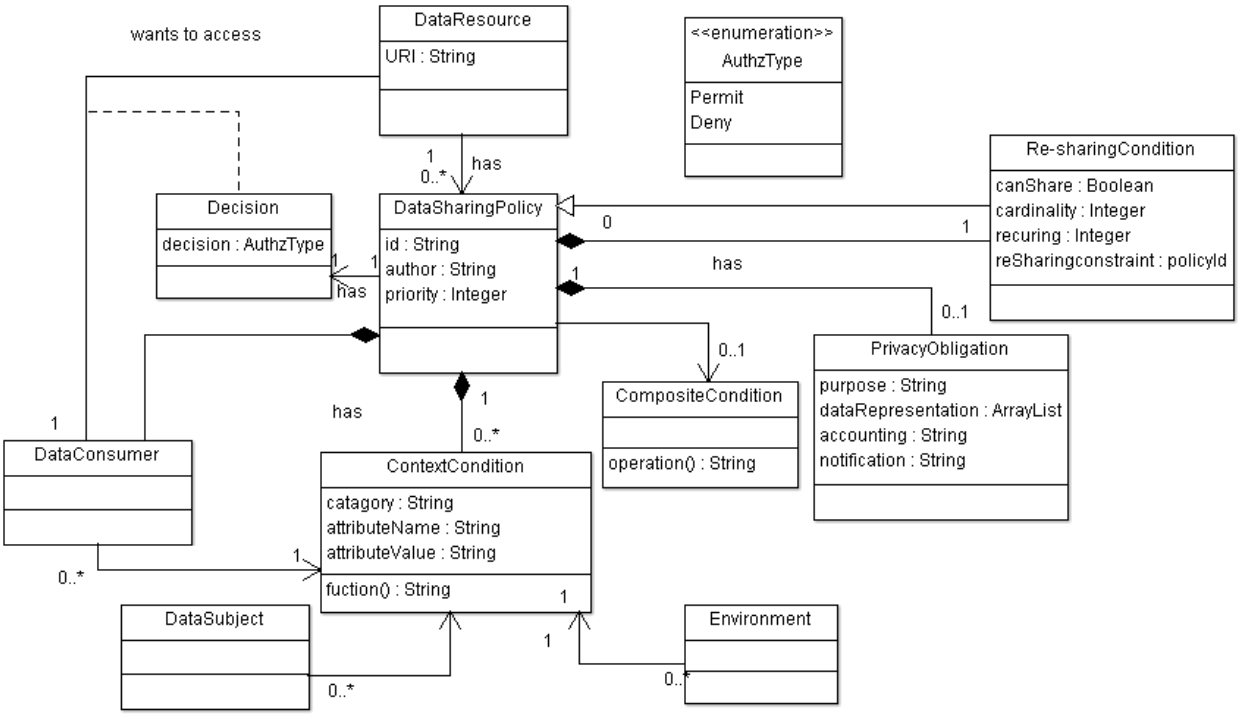

Figure 2. Meta-model of Data Sharing Policy.

data subject, data consumer and/or environment entity.

Privacy Obligation encode the data subject's instruction for which purpose this particular data can be used. PrivacyObligation also captures requirements related to data representation, notification and accounting.

The data custodian is responsible for enacting any enforceable privacy obligations defined by the data subject. The obligations include data representation, notification, and accounting. As obligations such as "purpose" cannot be directly enforced by the data custodian, a mechanism needs to be provided whereby, prior to accessing the data, the consumer must agree to use the data for the prescribed purpose before access is granted. Such agreement may be legally binding. Auditing and trust based scores can be used to help assure that this obligation is not violated by the data consumer.

ReSharingCondition is used to delegate "sharing rights" to the data consumer by the data subject and define the constraints on the re-sharing. Constraints are defined by another data sharing policy which is linked to the primary policy via a construct (*ReSharingConstraint*) in ReSharingCondition.

3.2. Concrete Syntax of Data Sharing Policy Language

Representational state transfer (REST) or RESTful approach has emerged as the most popular architecture to build web applications in recent times [9]. REST-compliant Web services allow requesting systems to access and manipulate textual representations of Web resources using a uniform and predefined set of stateless operations. As RESTful application is resource oriented, it is a natural way to reference data and access operations in a resource-oriented data sharing policy language. Utilizing resource orientation in access control also offers the potential for fast identification of policies that need to be evaluated for any access request. We use JSON [10] as the data format. A JSON object consists of a group of key-value pairs surrounded by curly brackets and objects can be nested. *Data Sharing Policy* uses the attributes *id, priority* and *author* for evaluation and conflict resolution.

Data Subject: is presented by its category and attributes. Category is used to link the attribute with any specific entity. For example, attribute "location" can be related to data subject or data consumer. Here, category is used to specify whether it is data subject or data consumer's location. *Data Consumer* and *Environment*

entities and their attributes are also defined in a similar way.

ContextCondition: is composed of one or more conditions defined by the attribute value of data subject, data consumer or environment entity.

Example: A data consumer's location must be equal to "XYZ office address" to get access.

```
{ "contextCondition" : [
"function"        :  "equal" {
"category"      :  "dataConsumer",
"attributeName" : "location",
"attributeValue": "Hawthorn" }]}
```

Privacy Obligation: is used to define the privacy preference of the data subject. We can used the JSON syntax like the *ContextCondition* to define privacy obligation.

Re-Sharing Condition: captures the requirements related sharing rights delegation and the constraints on re-sharing. We have shown how re-sharing conditions can be captured by concrete syntax in Case Studies section.

Example: Mr. Smith can re-share Alice's shared academic record but only with his two current colleagues. However colleagues cannot share them further.

4. Case Studies

In this section we present three scenarios to highlight context-aware, privacy preserving and re-sharing requirements in data sharing environment. We use our proposed data sharing policy language to encode data subject's requirements to demonstrate the applicability of our proposed language using the following three scenarios.

4.1. Context Aware Data Sharing Scenario

Alice is a university student. She lives alone. Because of the safety of her pet and other belongings, she has recently bought web cam service from an online company called webcam.com. They (webcam company) have installed a smart online web camera at her home. They have also provided her a user interface (at webcam.com - which is the data custodian in this case.) where she can access her (watch) web cam online. She is usually busy with her study. She wants to share (delegate access (only view)) her web camera access with her mother. However, she wants that her mother can only get access to the web camera when she (Alice) is not at home.

Data Sharing Policy: *"Alice's Mom can access Alice's web camera only when Alice is not at home."*

```
{ "dataSharingPolicy" : {
"id" : "P1",
"author" : "Alice",
"decision" :  permit",
"dataConsumer" : {
"attributeName" : "email",
"attributeValue" : "mom@example.com"},
"contextCondition" : {
"function  :  "not-equal",
"category" :  "dataSubject",
"attributeName" : "location",
"attributeValue": "home" }}
```

4.2. Privacy Preserving Data Sharing Scenario

Shaila is visiting Melbourne to see David. Shaila plans for a city tour to explore Melbourne. She shares her location with David so that he knows where she is. However, she does not want to share the exact location with David because of her privacy concern (e.g., she does not want to share which shop she is visiting).

Data Sharing Policy: *Ms. Shaila wants to share her street level address with David. She also wants notification when David access her location.*

```
{ dataSharingPolicy : {
"id" : "P2",
"author" : "Shaila",
"decision" : "permit",
"dataConsumer" : {
"attributeName" : "email",
"attributeValue":"david@example.com"},
"privacyObligation" :  {
"dataRepresentation" :
["granularity-location-street"],
"notification" : "shaila@gmail.com"}}
```

4.3. Re-sharing Scenario

Ms. Tanya and Mr. James work at a company called OfficeWork. She has prepared a project proposal for the client. She has shared the proposal with James for his feedback. Now she wants to delegate her "sharing

rights" to James so that James can re-share it with others to get more feedback. However, she wants to prevent James sharing the proposal with anyone outside of their company. She also wants that he (Mr. James) can only share it from 1st of March to 7th of March.

Data Sharing Policy: *Mr. James gets re-sharing rights on the project proposal. However, he can only share this document within his company colleagues and from 1st of March to 7th of June*

```
{dataSharingPolicy : {
"id" : "P3",
"decision" : "permit",
"dataConsumer" : {
"attributeName" : "email",
"attributValue":"james@oficewrk.com"},
"reSharingCondition" : {
"canShare" : "true",
"reSharingPolicyId" : "R2"}}}
{dataSharingPolicy: {
"id" : "R2",
"dataConsumer": {
"attributeName" : "email",
"attributValue :"*@officework.com"},
"compositeCondition" : {
"operation  :  "AND",
"contextCondition" : {
"function" : "greater-than-or-equal"
"category" :  "environment"
"attributeName" : "date"
"attributeValue : "01-06-2017"  },{
"function" : "less-than-or-equal"  {
"category" :  "environment",
"attributeName" : "date"},{
"attributeValue": "07-06-2017"}}}}
```

5. Data Sharing Access Control Model

In this section we show how the proposed data sharing policy language can be used in conjunction with an access control system. We discuss how the policy is bound to the resource and the protocol for handling access requests from a data consumer. We then briefly describe our implementation of the policy definition and access control system.

5.1. Policy Evaluation Algorithm

Existing access control (e.g., XACML[6]) and privacy policy languages (e.g., P3P [7], EPAL[8]) do not support all the requirements identified in the previous section . We have presented a brief description of the existing policy languages and their limitations in section 6. Thus, we propose a data sharing policy language that can be used by data subjects to define their data sharing requirements.

Algorithm 1: Policy Evaluation Algorithm

Input: Data Sharing Policy: Policy
Result: Access Decision: Decision

```
1  initialize decision as null, decision = null
2  authenticate the data consumer,
     authn ← authenticate (identifier)
3  if authenticated then
4  |   foreach context condition in Policy do
5  |   |   attValue ← reTrieveContext(cat, attName)
6  |   |   eval ← compareCtx(attValue,
   |   |     attValue,function)
7  |   |   if eval is true then
8  |   |   |   decision=permit;
9  |   |   else
10 |   |   |   decision=deny
11 |   |   end
12 |   end
13 |   purpose ← Policy[policyObligation][purpose]
14 |   checkPurpose(purpose, responseDataConsumer)
15 |   decision = permit;
16 else
17 |   decision = deny;
18 end
19 return decision
```

Re-sharing is a complex phenomena in our proposed policy framework. The data subject can delegate the "re-sharing" capability to the data consumer. If delegated then the data consumer can re-share it with other data consumer. However, the data consumer must comply with the primary data sharing policy and the re-sharing constraints defined by the data subject.

5.2. Data Sharing Protocol

Figure 3 shows the sequence of how the data sharing policy can be defined by the data subject and will be enforced in the data custodian.

First of all, Alice defines her data sharing policy related to her academic record hosted at ABC University. In this policy, Alice defines all her requirements related to this sharing. Then ABC University server generates a security token for Mr. Smith and sends the resource URL with security token to Mr. Smith (e.g., by email). When Mr. Smith clicks

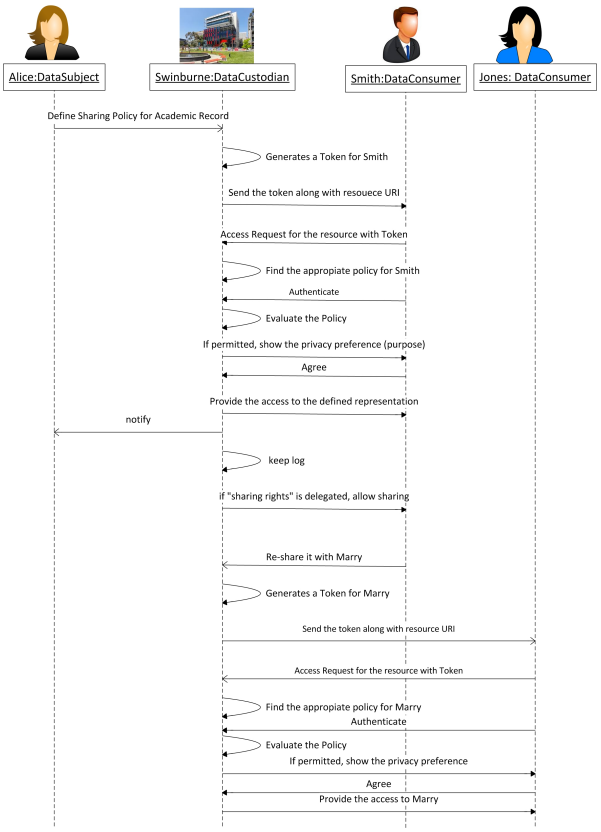

Figure 3. Sequence Diagram of the Data Sharing Protocol.

on the link ABC University server captures the request and sends the request to the Policy Engine hosted in the same server. Then data sharing policy is evaluated and if all the requirements (e.g., contextual condition, privacy obligation and re-sharing condition) are met then the "purpose" obligation confirmation is prompted to Mr. Smith. If Mr. Smith agrees then the ABC University server notifies Alice of the access as described in the policy and will provide the defined representation of data resource to Mr. Smith. Mr. Smith can also share the shared resource (as specified in the policy) with his colleagues (e.g., Ms. Jones from XYZ company).

5.3. Implementation

We have implemented a prototype to demonstrate how the data sharing policy can be defined and how that policy can be encoded in JSON. We have followed the scenario presented in section II. The implementation also followed sequence presented in the previous section.

We have used HTML, JavaScript, JSON, and PHP as the programming language and MySql as database. The data sharing policy is stored as .json file in the server and the resource-policy mapping is stored in the database. At first data subject login to the system then she clicks on the "share" button in the resource page. Then she will be provided with a policy defining page. She defines all her sharing requirements in that page and click on the "Generate Policy" button. It will generate a policy in JSON format and send an email with the resource URI and security token to the data consumer and store the resource and policy mapping in the database. When data consumer clicks on that link he will redirected to the data custodian The data custodian then has to authenticate using Federated Identity (e.g., OpenIdConnect, Australian Access Federation etc.) system. After successful authentication, the policy will be evaluated for context conditions then he will be asked to agree with the "purpose" of sharing. If all the conditions are fulfilled access will be granted. We have uploaded a demonstration video online to show the basic implementation[1].

5.4. Formal Modeling For Validation

Formal Modeling is an effective way to validate access control policies. Different researchers have used different mechanisms to validate the access control model and policies. Alloy [11] is one of the most popular model analyzing tool and has been used by several researchers to validate access control policies[12]. We have also used Alloy to model our proposed policy language for vitrification. We have formally defined our proposed policy language using the syntax of Alloy. shows Alloy generated meta-model of our proposed language. However, the validation of the model is in our immediate future plan.

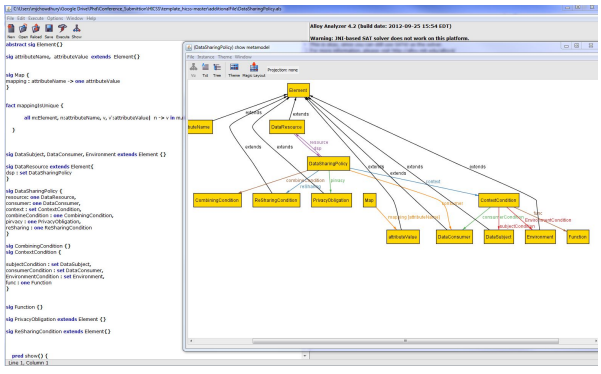

Figure 4. Formal Model Defined in Alloy

[1]https://www.youtube.com/watch?v=lRm0V_VDXjQ

Table 1. Comparison of Different Approaches with Our Proposed Approach

Criteria \ Approaches		[14]	[15]	[19]	[6]	[20]	[21]	[22]	[23]	Our Aprch
Context-Aware Access	Spatial	✓	×	✓	✓	✓	✓	×	✓	✓
	Temporal	×	✓	✓	✓	✓	✓	×	✓	✓
	Other	×	×	✓	✓	✓	✓	×	✓	✓
Privacy Obligation	Purpose	×	×	×	✓	✓	×	×	✓	✓
	Granularity	×	×	×	×	✓	×	×	×	✓
	Anonymization	×	×	×	×	×	×	×	×	✓
	Notification	×	×	×	✓	×	×	×	×	✓
	Accounting	×	×	×	✓	×	×	×	×	✓
Re-sharing Constraint	Access Delegation	×	×	×	✓	✓	×	✓	✓	✓
	Multiplicity	×	×	×	×	×	×	×	×	✓
	Multi-step	×	×	×	✓	×	×	×	×	✓
	Constraints	×	×	×	✓	×	×	×	×	✓
Subject-Centric	Can Define Policy	×	×	×	×	✓	×	×	✓	✓
	Resource Oriented	×	×	×	×	×	✓	×	×	✓

6. Related Work

Role Based Access Control (RBAC) model [13] has become the most widely used access control model. Several research efforts (e.g., [14], [15]) have been carried out to extend RBAC to incorporate context. The GEO-RBAC [14] model proposes the spatial extent (i.e., geographical location) of role in which the user is to be located for being enabled to play a role. Chandran et. al. [15] propose a location and time-based RBAC model, which uses temporal and location constraints for enabling and disabling of roles. However, all these approaches are limited by the number of contexts they can handle. Moreover, role based access control mechanism is not suitable for cross custodian environment because of the role-mapping problem [16]. To overcome this problem, researchers (e.g., [17],[18]) have extended the Attribute-based Access Control (ABAC) [19] approach to provide access control to resources in a context-aware manner.

Both types of approaches consider only the access control and do not cover individual's data privacy and re-sharing requirements. Moniruzzaman et. al. [20] proposed an access control model which considers delegation and privacy along with access control. Individual can define the policy to define their privacy and delegation preference. However, they have not considered context and re-sharing preferences.

XACML[6] is most popular policy implementation of Attribute Based Access Control model. It expresses access control policy using the attributes of the four types of entities, such as subject, resource, action and environment. It can also be used to capture context using dynamic attributes (e.g., location). Privacy construct is introduced in version 2 of XACML but it only covers "purpose" of the data sharing. Another limitation of XACML is that it is not fundamentally resource-oriented so mapping between resource and policies is not strait forward and identification of the relevant policy is slow.

P3P policy language [7] is used to specify data subject's privacy preferences regarding the handing of a particular data resource. However, it lacks the way of identifying a specific item of data or a specific data consumer. Furthermore, P3P cannot be automatically enforced. To overcome the shortcomings, IBM proposed the Enterprise Privacy Authorization Language (EPAL) [8]. It is intended to express an organization's privacy policy in such a way that it can be enforced by an access control system. EPAL has the element called "purpose" that makes permission conditional on the action being performed for some particular purpose. However, it also lacks the expression to define the granularity of the data resource. Hffmeyer et.al. [21] proposed an access control language for RESTful Services. Our approach follows a similar resource-centric approach, however the language in [21] is purely an access control model and does not provide any construct to capture privacy and re-sharing

requirements.

Modern authorization standards, like OAuth (RFC 6749 [22]), enable users to grant access to their data and process to third parties without disclosing the user's authentication data. User Managed Access (UMA [23]) is a protocol based on OAuth which enables the user to define policies. However, policy format and evaluation algorithm are not specified in UMA. Table 1 shows the comparison between our proposed solution with other existing solutions.

7. Conclusion and Future Work

Data sharing has become a common phenomenon in people's living in an increasingly digitised life. However, individuals (data subjects) are still not able to define their data sharing requirements in a comprehensive way, especially in a cross-domain (data custodian) environment. To overcome this limitation, this paper, we have first identified the various stakeholders and related concepts in a data sharing environment. Then we have proposed a data sharing policy language based on the identified stakeholders and concepts. Data subjects can use this proposed language to define their data sharing requirements with context conditions, privacy obligations, and re-sharing restrictions. We have carried out a case study with the policy language and defined a range of data sharing policies for data subjects, to validate its applicability. We have also developed a prototype implementation to support the definition of data sharing policies and their use and enforcement for runtime data sharing. In the future, we plan to further extend the implementation architecture to provide a central point from which individuals can define and enforce their data sharing policies for their data resources hosted in multiple data custodian systems. Furthermore, we will also provide a detail verification of our proposed model using Alloy [11].

References

[1] A. Poikola, K. Kuikkaniemi, and O. Kuittinen, "My data," 2014.

[2] G. Hornung, "A general data protection regulation for europe: Light and shade in the commission's draft of 25 january 2012," *SCRIPTed*, vol. 9, p. 64, 2012.

[3] A. Cavoukian, "Privacy by design [leading edge]," *IEEE Technology and Society Magazine*, vol. 31, no. 4, pp. 18–19, 2012.

[4] I. C. Office, "Data sharing code of practice." (2011).

[5] C. for Privacy and D. Protection, "Guidelines for sharing personal information." (2016).

[6] O. Standard, "extensible access control markup language (xacml) version 2.0," 2005.

[7] L. Cranor, M. Langheinrich, and M. Marchiori, "A p3p preference exchange language 1.0 (appel 1.0): W3c working draft 15 april 2002," *World Wide Web Consortium (W3C), URL: http://www. w3. org/TR/P3P-preferences*, 2002.

[8] C. Powers and M. Schunter, "Enterprise privacy authorization language (epal 1.2)," *W3C Member Submission*, vol. 10, 2003.

[9] J. Webber, S. Parastatidis, and I. Robinson, *REST in practice: Hypermedia and systems architecture*. " O'Reilly Media, Inc.", 2010.

[10] "Json (javascript object notation),." (2012).

[11] D. Jackson, I. Schechter, and H. Shlyahter, "Alcoa: the alloy constraint analyzer," in *Proceedings of the 22nd international conference on Software engineering*, pp. 730–733, ACM, 2000.

[12] M. Mankai and L. Logrippo, "Access control policies: Modeling and validation," in *5th NOTERE Conference (Nouvelles Technologies de la Répartition)*, pp. 85–91, 2005.

[13] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-based access control models," *Computer*, vol. 29, no. 2, pp. 38–47, 1996.

[14] M. L. Damiani, E. Bertino, B. Catania, and P. Perlasca, "Geo-rbac: a spatially aware rbac," *ACM Transactions on Information and System Security (TISSEC)*, vol. 10, no. 1, p. 2, 2007.

[15] S. M. Chandran and J. B. Joshi, "Lot-rbac: a location and time-based rbac model," in *International Conference on Web Information Systems Engineering*, pp. 361–375, Springer, 2005.

[16] L. Chen and J. Crampton, "Set covering problems in role-based access control," in *European Symposium on Research in Computer Security*, pp. 689–704, Springer, 2009.

[17] A. Corrad, R. Montanari, and D. Tibaldi, "Context-based access control management in ubiquitous environments," in *Network Computing and Applications, 2004.(NCA 2004). Proceedings. Third IEEE International Symposium on*, pp. 253–260, IEEE, 2004.

[18] R. Hulsebosch, A. H. Salden, M. S. Bargh, P. W. Ebben, and J. Reitsma, "Context sensitive access control," in *Proceedings of the tenth ACM symposium on Access control models and technologies*, pp. 111–119, ACM, 2005.

[19] L. Wang, D. Wijesekera, and S. Jajodia, "A logic-based framework for attribute based access control," in *Proceedings of the 2004 ACM workshop on Formal methods in security engineering*, pp. 45–55, ACM, 2004.

[20] M. Moniruzzaman and K. Barker, "Delegation of access rights in a privacy preserving access control model," in *Privacy, security and (PST), 2011 ninth annual international conference on*, pp. 124–133, IEEE, 2011.

[21] M. Hüffmeyer and U. Schreier, "Restacl: An access control language for restful services," in *Proceedings of the 2016 ACM International Workshop on Attribute Based Access Control*, pp. 58–67, ACM, 2016.

[22] D. Hardt, "The oauth 2.0 authorization framework," 2012.

[23] T. Hardjono, E. Maler, M. Machulak, and D. Catalano, "User-managed access (uma) profile of oauth 2.0," *Internet Engineering Task Force (IETF)*, 2015.