Information Integration and Analysis: A Semantic Approach to Privacy

Madan Oberoi*, Pramod Jagtap,
Anupam Joshi, Tim Finin
CSEE Department, UMBC
Baltimore, MD 21250

Lalana Kagal
CSAIL, Massachusetts Institute of Technology,
Cambridge, MA

Abstract— **The balance between privacy and security concerns is a hotly debated topic, especially as government (and private) entities are able to gather and analyze data from several disparate sources with ease. This ability to do large scale analytics of publicly accessible data leads to significant privacy concerns. In particular, for the government, there is the fear of a fishing expedition against individuals. The model in this paper describes a way to address these concerns in a multi-user and multi-database owner environment. The model provides an assurance system where database owners are able to test and audit the assurances given by users thereby increasing the trust in the system. The concept of segregating data used for processing from data needed for final end use and providing different levels of access to them through a mediator machine has been used. The audit component consisting of a justification mechanism increases the trust in the system.**

Keywords-Information; integration; analysis; semantic approach; security; privacy; trust; mediator machine; query manipulator; compliance screen; justification mechanism; audit control

I. INTRODUCTION AND MOTIVATION

In today's highly networked information infrastructure a huge amount of personal information is in the public domain, gathered by a variety of government and private entities. In the wake of increased incidents of terrorism at global level, various national security agencies have sought to access, integrate, and analyze more personal information. This in turn has led to privacy concerns. While the expectations of privacy by citizens vary with culture and country, it appears that often citizens are relatively more comfortable with commercial companies mining their personal information rather than law enforcement agencies collecting and mining this data across information sources.

The balance between privacy and national security concerns is very difficult to achieve, especially with the evolution of how the Law Enforcement Associates (LEA) and Counter-Intelligence agencies operate. Traditionally security agencies used to look for information about known or suspected individuals. Most national security agencies, in view of changed strategy involving preemptive identification of likely rogue elements, are relying more on tools like data mining and surveillance to identify patterns and inconsistencies that can indicate threats [1]. Security agencies have taken an array of data mining initiatives in this regard [2], [3], which has led to their ever increasing demands for access to more databases containing personal information. The tools used by these agencies for various kinds of pattern analysis typically need the entire data generally referred to as 'data dump' for analysis.

This, in many cases, results in conflicts with privacy policies of organizations and citizens, which may be willing to share specific information about a suspect for national security purposes, but are normally not amenable to providing the entire dump of data for a 'fishing operation'. From the point of view of national security agencies the lack of access to such data dumps denies them a chance of data analysis for preemptive identification of likely suspects, thereby denying them a head start in their fight against terrorism.

Efforts have been made to reconcile the two equally important, although sometimes conflicting, goals of national security and protecting citizens' expectation of privacy. This paper addresses this problem in order to explore a solution to balance both concerns, as having more security should not necessarily mean having less privacy [4]. We propose a model to solve this issue by using machine understandable and semantically rich descriptions of the a) data, b) policies governing access and privacy, and c) the query context.

A key element of our approach is distinguishing between the query originator, and the analysis routine which ingests the data and responds to the query. In most situations the national security and law enforcement agencies don't need to see the data dump; rather it is required for processing various queries, the result of which generally is specific data, which is the end-use data. The data being shared by organizations can therefore be classified in two categories i.e.

    i.   Data for processing the query
    ii.   Data for end use as result of the query

If the two forms can be distinguished and access to the first form, i.e. data for processing the query, be separated from the requesting (user) organization, the privacy concerns of database owner agencies can be addressed substantially. This can be done by using a mediator organization [5], where the dump data is accessed only by a trusted hardware and software platform, which would be capable of enforcing privacy policies of the data owner agencies on the information ultimately going out of this machine. This has been illustrated in Figure 1.
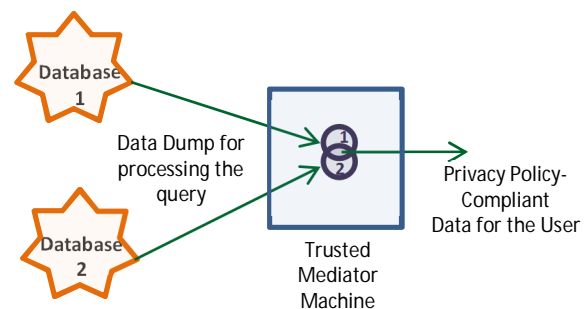

Figure 1: Separation of data for processing from end-use data.

Another element of our approach is articulating the context in which the query is made. The context of the query minimally includes who is asking for the information, and for what purpose. More generally, it includes an identification of the person or entity which initiated the query, their role in a hierarchy, the group(s) to which they belong, and the intended use of the information. In this sense, we capture the concepts associated with usage [6] and group based controls [7]. In order to address privacy concerns, organizations that collect personal data during their routine business prepare and publish privacy policies to assure their clients. These privacy policies determine the way, modalities, quantum, time period after which, conditions/situation under which, and with whom such personal information can be shared. We note that these policies are generally not machine interpretable or formal policies. However, by making them machine interpretable, we can reason over these policies, and the query context, to decide if the data can be shared.

An important feature of the conceptual model proposed in this paper is the system of automatic periodic audit to check whether the privacy policies were correctly enforced or not and to be able to throw up cases of exception for the follow up action. The concept of auditable policies is very important in cases where information is shared with 'after-access' obligations like some data may be shared with a condition to destroy it within X days from access. The audit component helps to assure the database owners that their privacy policies are being complied with. The conceptual model also includes providing justification for access decisions, which help the audit component determine whether the reasoning engine of model is arriving at correct inferences.

In this paper, we formally describe this conceptual model, and a realization of it using OWL (Web Ontology Language) [8] as our semantic description language for policies and query context, and Jena [9] as our reasoning infrastructure. We consider a hypothetical case in which the national security agency of a country is interested in doing preemptive identification of likely rogue elements by accessing dump data from four databases, including those of banks, passport office, telephone companies and immigration office. The case assumes that access to dump data by the national security agency is not allowed by the privacy policies of database owners, or legal and statutory obligations placed on them. In the proposed model the dump data is accessed by the trusted mediator system for processing the analytic query. The result of query, which is a collection of a limited number of records, is shared with the national security agency assuming that sharing even this limited information does not violate the policy of the database owner.

This paper has been organized in four main sections, besides this introduction. Section II deals with related work, while section III describes the proposed model. Section IV and V deal with implementation and Evaluation of this model.

II. RELATED WORK

The TAMI (Transparent Accountable Data-mining Initiative) project attempts to address issues of transparency, accountability in context of personal privacy by changing the perspective from controlling or preventing access to encouraging appropriate use of accessed data and inferring when data is misused by investigating the audit logs [10]. Our proposed work is closely related as it relies on logs to figure out whether obligations are met. However, unlike TAMI, our model does enforce privacy policies but does so on the end use data produced as a result of the query instead of the initial data dump required.

Kagal, Hanson and Weitzner [11] have discussed providing explanations associated with the derivation of a policy decision in the form of a list of reasons, called dependencies by them, using semantic web technologies. This kind of explanations will help the user as well as database owner agencies to understand how the results were obtained, thereby increasing trust in the policy decision and enforcement process. Our model will provide similar justifications about query decisions.

A lot of work has been done to develop machine interpretable policy frameworks [12], [13]. Rein (Rei and N3) [14] is a distributed framework for describing and reasoning over policies in the Semantic Web. It supports N3 rules [15], [16] for representing interconnections between policies and resources and uses the CWM forward-chaining reasoning engine [17], to provide distributed reasoning capability over policy networks. AIR [18] is a policy language that provides automated justification support by tracking dependencies during the reasoning process. It uses Truth Maintenance System [19] to track dependencies. Policies and data are represented in Turtle [20], whereas the reasoning engine is a production rule system [21] with additional features for improved reasoning efficiency such as goal direction. Rei and AIR consider rules defined over attributes of classes in the domain including users, resources, and the context. Though our initial prototype uses OWL to describe privacy policies, we plan to use AIR in the future to take advantage of its built-in justification feature.

Letouzey et al [22] have discussed existing security models by defining the security policy through logically distributing RDF data into SPARQL views and then defining dynamic security rules, depending on the context, regulating SPARQL access to views. Kagal and Pato [23] have explored the use of semantic privacy policies, justifications for data requests, and automated auditing to tackle the privacy concerns in sharing of sensitive data. Their architecture evaluates incoming queries against semantic policies and also provides a justification for permitting or denying access, which helps requesters formulate privacy-aware queries. Currently our conceptual model does not restrict the query language to be used, but we plan to use SPARQL for better integration with Semantic Web data sources.

III. MODEL STRUCTURE

In our proposed model, there are multiple users and multiple database owners. Each database 'D' has its own set of (privacy) policies D(P), which can be reduced to rules. Similarly each user, belonging to any of the user agencies, will have its own set of privileges U(Ø). These privileges, we are assuming, would depend upon the hierarchical position of user, his membership of various groups as well as the use for which information is being sought. A query (Q) made by a member of the user organization is therefore a tuple where

$$(Q_1, Q_2, Q_3, \ldots\ldots Q_n) \equiv (U, G, H)$$

and U is a set of uses for which response to the query is needed, G is a set of groups to which the query originator can belong and H is a set of hierarchy levels, in which the query originator can be placed:

$$U = \{U_1, U_2, U_3, \ldots U_x\}$$

$$G = \{G_1, G_2, G_3, \ldots G_y\}$$

$$H = \{H_1, H_2, H_3, \ldots H_z\}$$

An important component of this model is the trusted Mediator and Audit Control System. The mediator system performs the critical function of ensuring segregation of data used for processing the query and data being shared with the user and thereby enforcing privacy policies of database owners. The mediator machine typically has access to more data than made available to the user organizations and it also has access to the privacy policies associated with each database via its VOID description [9]. In this conceptual model we propose multi-layered checking of the compliance of privacy policies through the Mediator System, whose subcomponents include query manipulator, compliance checker and audit controller, which are discussed below.

A. Query Manipulator

The Query Manipulator performs some or all of the functions of

     i. splitting the query into various sub-queries addressed to different databases

    ii. rewriting the query to be able to deliver a privacy-policy compliant answer

   iii. negotiating the query, the answer to which apparently violates the privacy policies

The query manipulator of the mediator with the help of a federation engine [24] splits the query of a user into sub-queries, which are directed to different databases. At this point, the query manipulator, as the first layer, checks whether sub-queries meet the privacy policies of the individual databases, except the policies relating to data-dumps.

If the queries do not comply with the privacy policies of the databases, being accessed, query manipulator tries to rewrite the query so that it becomes compliant with the concerned privacy policies. The rewritten query would be supplied to user and if user decides to proceed with this query, query manipulator would execute this. An alternative approach to rewriting query is to negotiate. This, for example, would be applicable in those cases where privacy policies refer to contextual situation (e.g. declaration of security level red), whose existence may have to be certified by a competent authority (say above level H5 in the hierarchy) . In such a scenario, when the query manipulator returns the unexecuted query with an explanation about reasons for non-execution, user may provide it with additional data in terms of certification by the required level in the hierarchy. The functioning of query manipulator has been depicted in figure 2.

B. Compliance Checker

This is the most important part of the mediator system. Once the results of sub-queries have been received; processed and final result of query has been generated, the compliance checker examines the final result to check if it meets the privacy policies of all databases. At this stage the compliance screen also checks privacy policies relating to data dumps, which were not checked at the stage of query manipulator. There might also be cases in which data dumps, which may be prohibited by some privacy policies, would have been supplied at the stage of sub-queries, but if the final result is also in the form of data dump, the result is found to be non-compliant by the compliance checker. In case the final result is found to be non-compliant by the compliance screen, then instead of transmitting the non-compliant result, the query would be routed back to query manipulator for rewriting and negotiation.
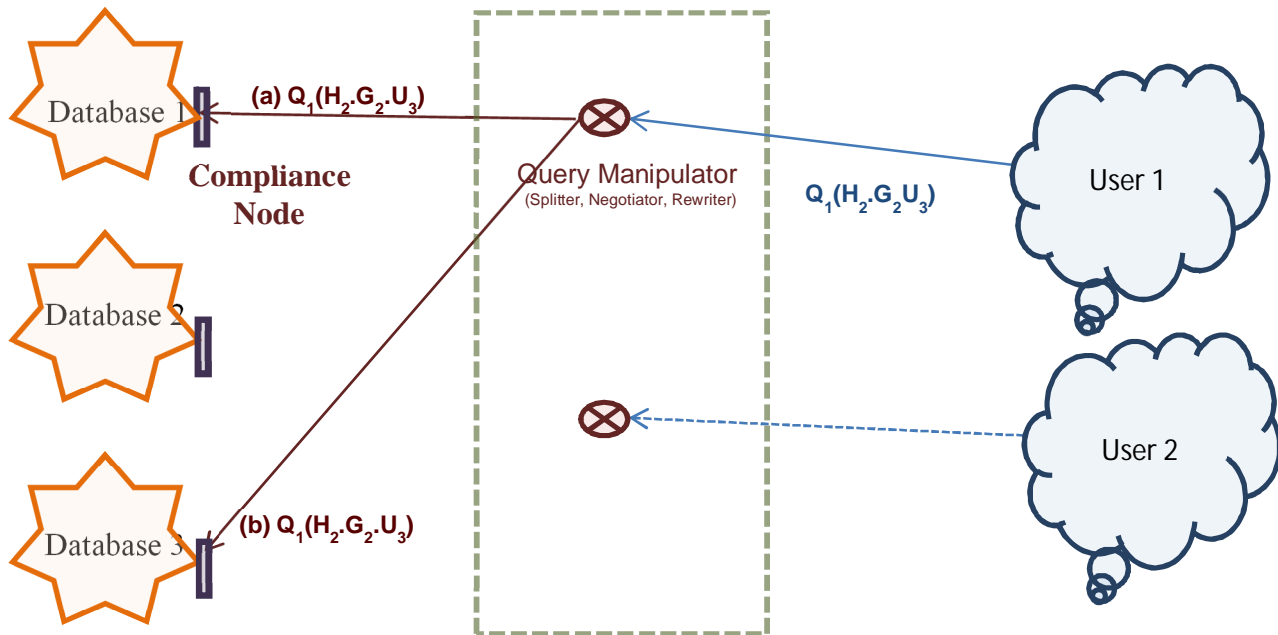

Figure 2: Query Manipulator in Mediator Organization

C. Audit Controller

The audit controller checks specifically for 'after-access' privacy obligations. For example, there may be restrictions on the uses for which information provided by the different databases can be used (e.g. not for tracking tax evasion) or there may be a condition about destruction of information after a stipulated time period. The information is provided on the assurance of the agencies that these 'after-access' obligations will be complied with. In order to ensure compliance, auditing is essential. The audit controller attempts to semi-automate this auditing process. It requires extensive logging of all use of information obtained from the mediator system. The logging system collects justifications for all actions performed on this information that provide extensive contextual information about each action including the person performing it, the purpose for the action, and the result. Since the original obligations are machine understandable, the audit controller reasons over these logs and justifications for possible non-compliance. On the basis of such an audit, periodic reports are sent to all database owners, mediator organizations and user agencies. This has been depicted in figure 3. It is proposed that this audit controller be implemented through a trusted third party. In most governments, such independent audit agencies exist and have statutory or constitutional authority.

IV. IMPLEMENTATION

We have taken an example of a national security agency to illustrate implementation of the above discussed model and concept. In the wake of increasing terrorist incidents this fictitious national security agency feels the need for "pattern-based" data mining. In order to do this they have recognized the need to shift from 'need-to-know' paradigm to 'need-to-share' amongst various law enforcement, tax and national security agencies, as well as their need to get data from private entities such as telecom providers, ISPs and banks. In order to do this they have proposed a distributed grid/services type framework where databases of various owner organizations and user agencies would be connected. The databases to be connected include those belonging to departments of immigration, telecom, bank, passport, tax, police, vehicle registration, driving licenses, airline and railway transport etc. However, privacy constraints prevent unfetterd gathering of information. The model discussed in previous sections addresses privacy issues in such multi-user, multi-owner situations like these, where there is a clear trust deficit with regard to capabilities as well as intentions of user agencies in complying with privacy policies of database owners (figure 3). Our current prototype implementation mainly evaluates the concept of providing more access to the trusted mediator machine; components related to compliance screen; and providing justifications draw by reasoning engine. We do not focus on the query manipulator since that is a well-studied area in database and semantic web systems.

The proposed model takes into account all sets of privacy policies of these database owners (shown as P1, P2, P3 and P4 in above diagram) and sets of privileges (depicted as Ø1, Ø2, and Ø3), which are dependent on hierarchical position level (H1, H2, H3, H4 and H5), Group Membership (G1, G2, G3 and G4), and Use (U1, U2, U3 and U4), for each of the user agency. The model includes privacy control module at compliance node to ensure that query results are compliant with privacy policies of database owners. In the proposed model it has been assumed that the national audit organization would implement the audit control structure to assure various database owners.
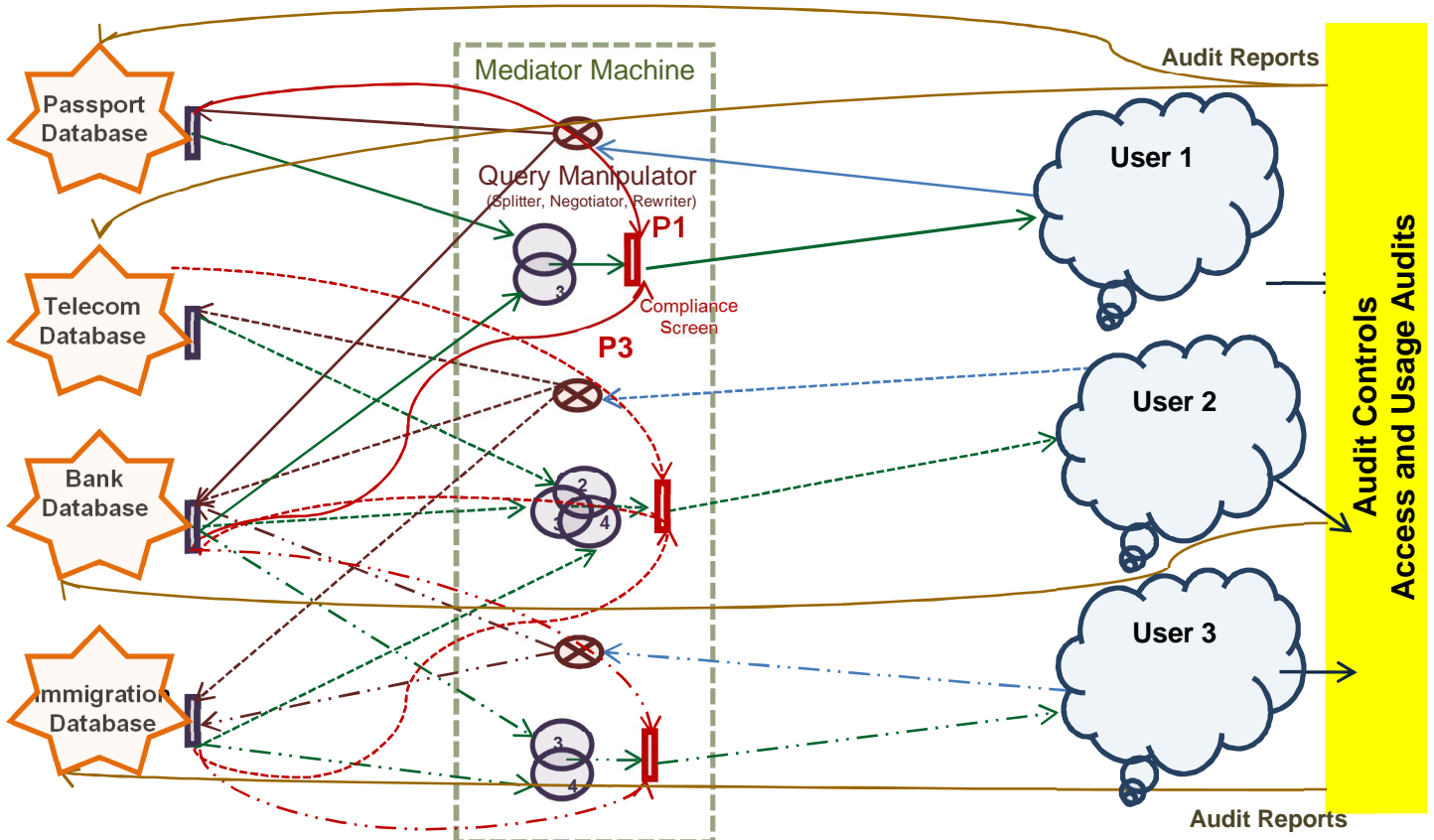

Figure 3: Model depicting Mediator Organization and Audit Structure

The compliance checker, which is the privacy control module, consists of

    i.   Ontology: defines various entities in order to make access control decisions,

    ii.   Information associated with the entity requesting queries, such as hierarchical position level, group membership and uses,

    iii.   Privacy policies defined by database owners, and

    iv.   Reasoning engine to perform reasoning to ensure compliance with privacy policies.

Our ontology describes the notion of hierarchical position level, group, and use. We have adopted description logics (DL), specifically OWL, and associated inferring mechanisms to develop the model and policies. The requester information consists of his position in the hierarchy, his group membership and use for which information is being sought. In our system this information is represented in N3 [15] as shown in figure 4. *Nat* is the namespace of our ontology while *foaf* is the FOAF vocabulary [25], which allows users to describe personal information about themselves and their relationships. The *Nat* ontology defines various properties such as *'belongs_to_hierarchyLevel'*, *'has_designation'* and *'belongs_to_group'* that can be used to represent the requester details. This information is used to determine whether the requester has the permission to access the query result based on database owner's privacy policies. The reasoning engine performs reasoning over this information and privacy policies. Our system uses the Jena Semantic Web framework [26] [27] for reasoning over the context data and the policies. Jena inference system allows the support of various inference engines or reasoners. These reasoners are used to infer additional facts from the existing knowledge base coupled with ontology and rules. The instance of such reasoner with a ruleset can be bound to a data model and used to answer queries about the resulting inference model. In our system, the reasoning engine uses the *Nat* ontology and the FOAF ontology to represent the requester information, and privacy policies represented in the Jena rule language to generate an inference model. This inference model is used to decide whether the information can be released to requester.

```
:Tom a foaf:Person ;
    a Nat:Requester;
    foaf:name "Tom" ;
    Nat:has_designation Nat:Additional_Secretary ;
    Nat:belongs_to_hierarchyLevel Nat:H1;
    Nat:belongs_to_group (Nat:G1 Nat:G2) .
```

Figure 4: User information represented in N3.

The privacy policies are rules that describe how a database owner wants to share information; with whom, and for which uses. For instance, the passport database owner can have privacy policies, which may have the restrictive condition about following types of information:

- Data Dumps– these cannot be directly given to any user,
- Information which has complete details to fix individual identity, i.e. Personally Identifiable

Information (PII), This can be released only if the request comes from persons at hierarchy level H2 or above, from groups G2 or G4 and for use U1 or U3.

This policy is shown in figure 5 and is represented in Jena syntax. Similarly, the immigration database owner can specify privacy policy by putting restrictions on following types of information:

- Data dumps – these cannot be directly given to any user,
- PII can be shared only if the request comes from persons at hierarchy level H2 and for use U1.

Figure 6 shows the Jena syntax representation for this policy. For the purpose of implementation and evaluation of this model, we have assumed the definition of data dump as a compilation of more than 5 records

V.   EVALUATION

The goal of evaluation was to see if the system satisfies the basic criterion of allowing the query result information to be shared with privileged users even though the users might not be permitted to access the intermediate data dumps required for query processing. To perform system evaluation, we designed use cases with sample user information and various privacy policies. Each of these use cases has either a different requested resource or different requester. The results of these use cases were initially inferred manually and then compared with actual system results with same settings. The system behaved as expected by allowing information access to privileged users and denying access to illegal users as per privacy policies, even when the trusted mediator was permitted to access to more information.

We developed privacy policies for passport, immigration, bank and telephone databases. In this system, we assumed that the privacy policies of passport database and immigration database do not allow disclosure of dump data. Though access to dump data was allowed to the trusted mediator, the compliance checker ensured that dump data was not available to the end user agencies. Similarly in this model, we assumed that the privacy policy of passport database does not allow disclosure of information, which has complete details to fix individual identity, i.e. personally identifiable information (PII) information such as records containing all the information on Passport Number, Date and Place of Issue of Passport, Name, Date of Birth, Address, Place of Birth etc. The policy allows disclosure of PII only if the request comes from persons at hierarchy level H2 or above, from groups G2 or G4 and for use U1 or U3. Similarly we assumed that the privacy policy of the immigration database does not allow disclosure of an individual's PII, which can be shared only if the request comes from persons at hierarchy level H2 and for use U1. These policies are represented in Jena rules as shown in figure 5. We considered a use case with requester "Tom", as shown in figure 4, who has the designation of "Additional Secretary", which belongs to hierarchy level H1 and he belongs to groups G1 and G2. 'Tom' asks for the list of passport holders from a specific area that have applied for immigration in the last year. This requires a data dump of the passport database, which is not permitted for Tom but the mediator system is allowed to proceed with the query processing. The result of the query is a

small set of records and is found to be compliant by the compliance checker. But in other cases, like for a different use, or a request for data dump, the system denied information sharing with "Tom".

```
[Rule1_1:
        (?requester a Nat:Requester)
        (ex:output ex:recordCount ?records)
        lessThan(?records, 5)
        ->
        (?requester :isAccessingDump "False")
]
[Rule2_2_1:
        (ex:output ex:has_passportNumber "True")
        (ex:output ex:has_dateOfBirth "True")
        (ex:output ex:has_placeOfBirth "True")
        (ex:output ex:has_name "True") (ex:output
        ex:has_address "True") (ex:output
        ex:has_passport_issue_details "True")
        ->
        (?requester :hasAllFields "True")
]
[Rule2_2_3:
        (?requester Nat:belongs_to_group ?groups)
        listContains(?groups, Nat:G2)
        ->
        (?requester :belongsToAllowedGroup "True")
]
[Rule2_2_4:
        (?requester Nat:belongs_to_group ?groupList)
        listContains(?groupList, Nat:G4)
        ->
        (?requester :belongsToAllowedGroup "True")
]
[Rule2_2_5:
        (?requester Nat:has_use ?uses) equal(?uses,
        Nat:U1)
        ->
        (?requester :hasAllowedUses "True")
]
[Rule2_2_6:
        (?requester Nat:has_use ?uses) equal(?uses,
        Nat:U3)
        ->
        (?requester :hasAllowedUses "True")
]
[RulePassportDatabse:
        (?requester ex:isRequester "True") (?requester
        :isAccessingDump "False") (?requester
        :hasAllFields "True") (?requester
        Nat:belongs_to_hierarchyLevel
        ?requesterLevel) (?requesterLevel
        Nat:higherLevel_than Nat:C) (?requester
        :belongsToAllowedGroup "True") (?requester
        :hasAllowedUses "True")
        ->
        (?requester ex:access ex:permitted)
]
```

Figure 5: Jena rule corresponding to privacy policy specified by Passport database owner.

When another requester "Harry" having designation "Director", which belongs to hierarchy level H3 and belonging to group G1, requested specific user's passport information for use U1, the system rejected request as requester did not have required hierarchy level

```
[RuleImmigrationDatabase:
        (?requester ex:isRequester "True") (?requester
          :isAccessingDump "False")  (?requester
        ex:immigrationPersonalQuery ?queryValue)
          equal(?queryValue, "True") (?requester
        Nat:belongs_to_hierarchyLevel ?requesterLevel)
        (?requesterLevel Nat:higherLevel_than Nat:C)
              (?requester Nat:has_use ?uses) equal(?
                          ses, Nat:U
        )
->
(?requester ex:access ex:permitted)
]
```

Figure 6: Jena rule for privacy policy enforeced on Immigration Database

In all use cases, the reasoning engine used requester information and privacy policies to decide whether information sharing should be permitted. Our system provides justification for each of these decisions using Jena's justification mechanism. The justification mechanism is supported by derivation logging in Jena. These records are used to determine how an inferred triple is derived from a set of source triple and a reasoner. In other words, it enables tracing all the rules which lead to a given inferred triple. This can be used to audit the correctness of inferences presented by reasoning engine.

VI. CONCLUSION

The model described above addresses the privacy concerns in a multi-user and multi-database owner environment. The model shows a way to meet demands of access to various databases by national security and law enforcement agencies without sacrificing the privacy concerns. It provides an assurance model where by database owners are able to trust the assurances of users by making use of various audit components of the model.

The model describes the key concept of segregating access to data used for processing from access to data needed for final end use. The model uses the concepts of a mediator machine capable of reading machine-interpretable privacy policies and enforcing them through critical components like, Query manipulator, Compliance Screen and reasoning engine. The model also utilized the audit component consisting of a justification mechanism to check the correctness of inferences drawn by machine relating to access decisions.

REFERENCES

[1] U.S. General Accounting Office, "Data Mining: Federal Efforts Cover a Wide Range of Uses", (GAO-04-548), May 2004, at 3, 27-64, http://www.gao.gov/new.items/d04548.pdf.

[2] Department of Homeland Security, "Report to Congress on the Impact of Data Mining Technologies on Privacy and Civil Liberties", 7 (2007), http://www.dhs.gov/xlibrary/assets/privacy/privacy_rpt_datamining_2007.pdf.

[3]   Department of Homeland Security, "Report to Congress on the Impact of Data Mining Technologies on Privacy and Civil Liberties" 8 (2006), http://www.dhs.gov/xlibrary/assets/privacy/privacy_data_%20mining_%20report.pdf;

[4]   Conference Report Cantigny Conference Series, "Counterterrorism Technology and Privacy", McCormick Tribune Foundation, 2005, http://www.mccormickfoundation.org/publications/counterterrorism.pdf

[5]   Gio Wiederhold, "Mediators in the Architecture of Future Information Systems", IEEE Computer, March 1992, pages 38-49.

[6]   Jaehong Park, Ravi Sandhu "The UCON$_{ABC}$ usage control model", ACM Transactions on Information and System Security (TISSEC) Volume 7 Issue 1, February 2004ACM New York, NY, USA

[7]   R. Krishnan, R. Sandhu, J. Niu, and W. H. Winsborough. "A conceptual framework for group-centric secure information sharing". In ASIACCS '09: Proceedings of the 4th International Symposium on Information,Computer, and Communications Security, pages 384–387, New York, NY, USA, 2009. ACM.

[8]   W3C, "OWL Web Ontology Language", February 2004, http://www.w3.org/TR/owl-features/

[9]   "VOID – Vocabulary of Interlinked Datasets", http://semanticweb.org/wiki/VoiD

[10]   Daniel J. Weitzner, Harold Abelson, Tim Berners-Lee, Chris Hanson, James Hendler, Lalana Kagal, Deborah L. McGuinness, Gerald Jay Sussman, K. Krasnow Waterman, "Transparent Accountable Data Mining: New Strategies for Privacy Protection", MIT CSAIL Technical Report-2006-007, http://www.w3.org/2006/01/tami-privacy-strategies-aaai.pdf

[11]   Lalana Kagal, Chris Hanson, Daniel Weitzner, "Using Dependency Tracking to Provide Explanations for Policy Management", IEEE Policy: Workshop on Policies for Distributed Systems and Networks, June 2008

[12]   Tim Moses. eXtensible Access Control Markup Language TC v2.0 (XACML), February 2005.

[13]   Sushil Jajodia, Pierangela Samarati, V. S. Subrahmanian, and Elisa Bertino. "A unified framework for enforcing multiple access control policies" In Proceedings of ACM SIGMOD International Conference on Management of Data, pages 474–485. ACM Press, 1997.

[14]   Lalna Kagal and Tim Berners-lee. "Rein : Where policies meet rules in the semantic web", Technical report, Laboratory, Massachusetts Institute of Technology, 2005.

[15]   TimBerners-Lee and Dan Connolly, "Notation3 (N3): A readable RDF syntax", Technical report, 2008.

[16]   Tim Berners-Lee, Dan Connolly, Eric Prud'hommeaux, and Yosi Scharf, "Experience with n3 rules", In Rule Languages for Interoperability, 2005.

[17]   Tim Berners-Lee, "Cwm - a general purpose data processor for the semantic web".

[18]   Lalana Kagal, Chris Hanson, and Daniel Weitzner, "Using dependency tracking to provide explanations for policy management", In Proc. IEEE Workshop on Policies for Distributed Systems and Networks, pages 54–61, Washington, DC, 2008. IEEE Computer Society.

[19]   Jon Doyle, "Truth maintenance systems for problem solving", Technical report, Cambridge, MA, USA, 1978.

[20]   D. Beckett, "Turtle - Terse RDF Triple Language", Technical report, 2007.

[21]   D. A. Waterman and F. Hayes-Roth, editors, "Pattern-Directed Inference Systems", 1978.

[22]   Gabillon, A. Letouzey, L. Univ. de la Polynesie Francaise, Faaa, French Polynesia, "A View Based Access Control Model for SPARQL", Network and System Security (NSS), 2010 4th International Conference, September 2010, Melbourne.

[23]   Lalana Kagal_ and Joe Pato, "Preserving Privacy Based on Semantic Policy Tools", IEEE Security & Privacy Magazine Special Issue on: "Privacy-Preserving Sharing of Sensitive Information" August 2010, http://dig.csail.mit.edu/2010/Papers/IEEE-SP/db-privacy.pdf

[24]   Mathew Cherian, "A Semantic Data Federation Engine", MIT Masters Thesis Jan 2011

[25]   "The Friend Of A Friend (FOAF) Project ",http://www.foaf-project.org/

[26]   "Jena – Semantic Web Framework for Java", http://jena.sourceforge.net/

[27]   Carroll et al, "Jena: implementing the semantic web recommendations", ACM, pages 74-83, 2004