Semantic Enforcement of Privacy Protection Policies via the Combination of Ontologies and Rules

Yuh-Jong Hu, Hong-Yi Guo, and Guang-De Lin
Department of Computer Science,
National Chengchi University
Wen-Shan District Area, Taipei, Taiwan
{hu,g9607,g9610}@cs.nccu.edu.tw

Abstract

We propose that the semantic formal model for P3P and EPAL-based privacy protection policies can be enforced and expressed as a variety of ontologies and rules (ontologies+rules) combinations, such as DLP, SWRL, AL-log, DL-log, DL+log, and MKNF, etc. Based on P3P and EPAL's original expressions and their dictionaries, several ontologies+rules semantic enforcement of privacy protection policies will be proposed in this study that can be compared with existing others. Furthermore, we express privacy protection management policies as a set of ontology statements, rules, and facts for both information disclosure and rights delegation using one of the above ontologies+rules combinations for two specific use case scenarios. When verifying P3P/EPAL formal semantics, we exploit which ontologies+rules combination will be a feasible information disclosure control scenario under certain conditions. We hope that this study might shed some light on the study of future general information disclosure and rights delegation controlled on the open Web environment.

1. Introduction

When we consider the information disclosure problem, it is highly relevant to the privacy protection issue on the Web because both of them have to achieve the objective of information disclosure at the right time for the right persons (or agents) with the right purposes [10]. People always enforce very strict information access control policies in the centralized system where all of the users are already registered with their true identities and profile information. Once a user's account is granted for accessing the system resources, he/she should show his/her own pre-authorized user name and password to execute the intended software or to access sensitive information within this system.

However, this access control scenario cannot be easily enforced and implemented on the open Web where there are so many websites within it for users to randomly surf and search for their intended information [21]. In fact, it is still a big challenge to deal with the design and implementation of access control (or more specific privacy protection) policies and language on the open Web [26] [27].

It is impossible to compel a user to disclose his/her own profile information unless a particular website has enough incentives for the user to disclose his/her personal profile information. Furthermore, the disclosed user profiles might not be truly trusted and authenticated information because the user is afraid of his/her personal digital traces might be collected and analyzed later on for possible personal privacy invasion. The reason a user does not intend to disclose his/her personal profile for a website is that he/she is unaware how the collected personal profile and digitally traced information will be used. Even the website provides explicit usage statements that claims it will comply with existing legal privacy protection regulations. It is still very difficult for a user to justify whether the usage of collected profile data and digital traces are honestly compliant with the privacy protection statements indicated on that particular website [2].

The Platform for Privacy Preferences (P3P) is a privacy markup language for a web server to easily annotate a server's intentions on his collection of selective personal information usage options. Thus the P3P enables website to express its privacy practices in a standard format that can be easily and automatically retrieved and interpreted by user agents. P3P user agents will allow users to be informed of site practices (in both machine- and human-readable formats) and to automate decision-making based on these practices when appropriate [7]. On the other hand, using a P3P Preference Exchange Language (APPEL), a user can express his/her preferences in a set of preference-rules (called a ruleset), which can then be used by his/her user agent to make automated or semi-automated decisions regarding the acceptability of machine-readable privacy policies from P3P enabled Web sites [8].

Unfortunately, XML-based markup languages, such as P3P and APPEL, do not have the expressive power to model and enforce the semantics of a person's privacy protection intentions from both client and server sides [17] [29]. Therefore, we cannot simply use XML-based P3P to specify our privacy protection policies and ensure that these policies can be automatically verified to comply with the underlying legal regulations from a semantics unambiguous perspective. Obviously, we need a higher formal semantics layer laid on the P3P/APPEL to ensure all of the semantic clearness for policy compliance checking, which has been a very important research area of trust management for policy making [4].

If a website is trustworthy, we might allow it to freely collect our personal profiles and digitally traced information because we have confidence that this website will abide by the legalized information sharing and disclosure policies under the law per se to respect our intentions and options on the data usage. However, the problem is whether a trusted website can really be aware of the usage options and purposes for the collected information coming from a tremendous amount of different users on their selective options of personal profiles and digital traces. Another issue is whether we allow our agents to enforce the privacy protection principles unambiguously without our direct intervention. If possible, we might need to delegate privacy protection compliance checking services to one of the trustworthy web site's agents to ensure our benefits.

2 Related Studies

The P3P/APPEL privacy protection mechanisms were proposed to enable easy collection of data user (or data consumer) and date subject's (or data owner) usage purposes and conditions under a client server model [7] [8]. However, P3P can not support any semantics level representation and enforcement of privacy protection policies because P3P/APPEL expressions were based on XML syntax only. Similarly, the E-P3P (or later EPAL) was based on previous Flexible Authorization Framework (FAF) [28] [13] that was proposed to express and enforce the enterprise's privacy protection policies on the Web [15] [16]. The EPAL was using a logic program (LP) model to indicate the data usage purposes for a particular role under certain conditions. But the semantic representation and enforcement from the logic program model are still far from satisfactory from semantic representation and enforcement viewpoints.

While EPAL and XACML are proposed as privacy policy languages, they are very similar in both structure and in concept but the differences between these two languages are significant. Anderson argued that enterprises should choose XACML as a privacy policy language because the functionality of XACML 2.0 is a superset of EPAL 1.2 [1]. In order to ascertain all of information disclosure actions will abide by its privacy protection regulations; current P3P/EPAL privacy protection mechanisms were implemented and embedded into the relational database, such as Oracle Virtual Private Database (VPD) [17]. But it is not easy to exchange and share personal data and digital traces from heterogeneous data sources under VPD architecture. Certainly it is not easy to exercise the auditing and policy compliance checking based on the current P3P/EPAL design and implementation mechanism [2]. We need to have a more general and powerful semantic representation and enforcement of privacy protection framework to deal with the possible challenges that cannot be resolved by P3P/EPAL alone.

Previous semantic policy languages for security or privacy control were proposed either using description logic ontologies or logic program rules alone. For example, KAoS, Rei, and Ponder policy languages were based on ontology representation and reasoning only, so they were pretty limited with respect to their policy representation and enforcement [14] [25]. Similarly, the rule-based policy representation faced the same limitations on policy expression and enforcement [3] [6] [26]. In general, these policy languages have less expressive power compared with our ontologies+rules combination for policy representation and enforcement.

3 Privacy Protection on the Web

The original idea of WWW (or Web) is to promote the effective sharing and exchange of information among agents and people. After several years of development, the primary objective of this concept has been achieved to some extent. However, we expect information collectors provide the information disclosure option services for us as respecting our basic human rights while they are collecting and sharing our personal profiles and digitally traced information among themselves. In fact, this privacy protection consideration has become one of the most important emerging research issues in Web development.

3.1 Privacy Protection on Web 1.0 and Web 2.0

The Web 1.0 and Web 2.0 privacy protection problem is the primary focus for most of the current privacy protection languages, such as P3P/EPAL and XACML [1] [7] [15]. However, the systems enforced with these privacy protection languages cannot deal with the digital trace protection and disclosure issues because users' digital traces are usually stored as unstructured text-based weblog files. Furthermore, to achieve information sharing and exchange objectives on the current Web 1.0 and Web 2.0, we need to collect most of this information from multiple relational databases in the deep web. Therefore, the information disclosure policies and mechanisms for satisfying privacy protection principles are usually embedded into the relational database systems. Certainly it is not easy to have either information sharing or information disclosure actions across multiple relational databases from these heterogeneous database schema.

3.2 Privacy Protection on Web 3.0

If we successfully migrate from Web 1.0 and Web 2.0 to Web 3.0 (Semantic Web), then all of users' profiles and digitally traced information will be annotated via ontology-based markup language, such as RDF(S) or OWL. We do not know whether this semantic web evolution process will be a benefit or a detriment on the realizing of privacy protection. From the pro side, because all of the information will be modeled and marked up by a well-defined semantic web ontology structure we can easily apply similar techniques for the expression and enforcement of privacy protection policies. On the con side, if the privacy protection system was not perfectly designed and implemented, certainly this system would be much easier for any privacy violators to challenge our protection policies by using pre-existing highly semantically connected information to inference (or reason) where are the possible weak links to attack. This is a two-edged sword scenario when we introduce the semantic techniques into both information modeling and privacy protection policy on Web 3.0.

4 Ontologies+Rules for Privacy Protection Policy

We usually classify ontologies+rules combination as two approaches: homogeneous integration and hybrid combination [18]. In a homogeneous integration, ontologies will be the main body of concept for information structure, where DLP is the most restricted one for this approach [11]. All of the major terms and representations for privacy protection will be declared and defined in ontologies and later move to the rules for further inferencing processes, such as SWRL [12]. So the knowledge flow is uni-directional in a homogeneous integration of ontologies+rules. Here rules can be regarded as an added-on component to the ontologies component to enhance/extend the expression limitations of ontologies.

In a hybrid combination, the ontologies module is represented as OWL or RDF(S) and it sits side by side with the rules module represented as RIF to enforce the knowledge representation and integration on the well-known semantic web layer cake [1]. There are several possible hybrid ontologies+rule combinations, such as *AL-log*, *DL-log*, and *DL+log*, to consider as a policy language for the representation and enforcement of privacy protection policies [9] [23] [24]. Under hybrid ontologies+rules combination, some of the terms in privacy protection policies will not be explicitly declared or defined in ontologies but they will be declared as predicates in each rule. Therefore the knowledge flow between ontologies and rules might be bi-directional to re-enforce ontologies and rules expressive power of each other. At this moment, it is unclear which homogeneous/hybrid ontologies+rules combinations to use as an ideal representation and enforcement of privacy protection system policies. This still needs further study.

Another issue is that most of the current privacy protection systems with their policies can be expressed and enforced only as positive permission but no negative permission (or deny) on the rule's conclusion for each information disclosure request. Similarly, people do not allow weak (or strong) negation premises on each privacy protection policy. All of these constraints are due to the lack of negation as failure (NAF) assumptions for ontologies that certainly restrict wide information dissemination and disclosure capacity on privacy protection. In fact, how to merge open world assumption (OWA) from the ontology side with closed world assumption (CWA) from the rule side on the ontologies+rules integration is also one of the emerging critical research issues when we combine ontologies with rules together. Furthermore, we might face ontologies merging and rule composition challenges when we integrate the information cross heterogeneous multiple domains that might induce a dilemma for a global inconsistent ontologies+rules protection polices from each collected local consistent ontologies+rules protection policies [5].

4.1 Ontologies for Privacy Protection Policies

We proposed three types of ontology in the *DL+log*-based ontologies+rules combination for the semantic enforcement of privacy protection policies, e.g., data user ontology, data type ontology, and purpose ontology. More detailed structures with their associated class and property hierarchies are shown as the followings:

1. The structure of data user ontologies for both class and property hierarchies are proposed to categorize the type of users and with their memberships corresponding to an organization (see Figure 1).

2. The data type ontologies to describe both the hierarchies of class and property for personal profiles and

[1] See W3C Semantic Web Activity for the latest "layercake" diagram at http://www.w3.org/2001/sw/.

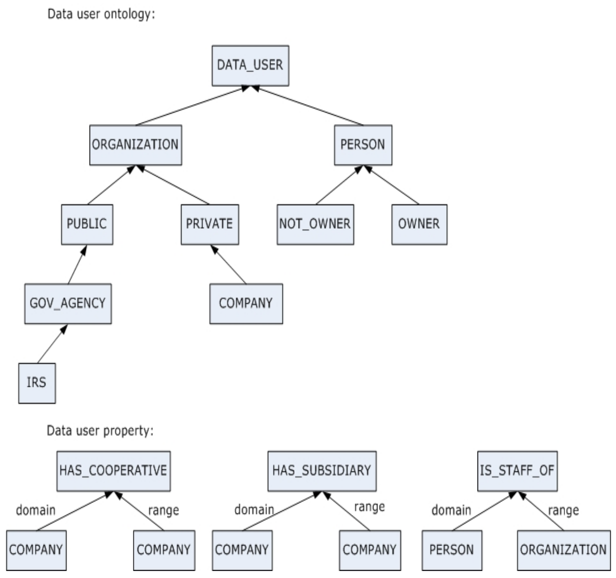


Figure 1. A data user hierarchy to classify the data user class hierarchy

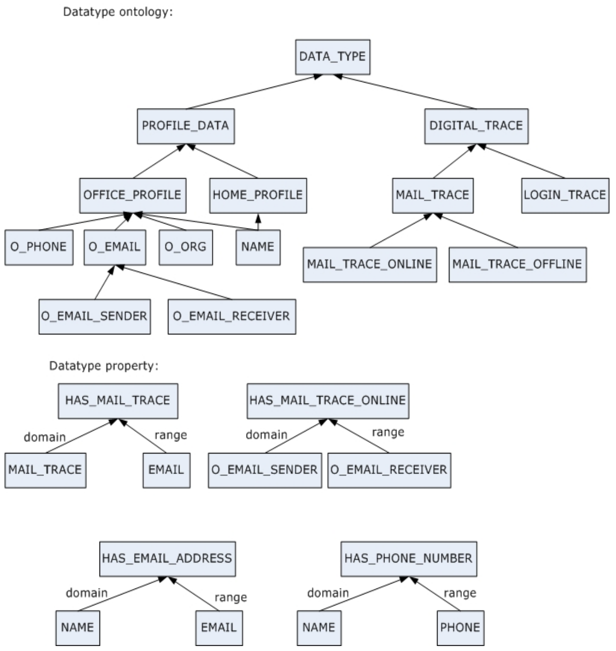


Figure 2. A class hierarchy classification for both personal profiles and digital traces

digital traces can be shown as Figure 2.

3. The purpose ontology to describe the intention of data user to use a particular type of data can be shown as Figure 3.

4.2 Two Scenarios for Privacy Protection of Mail Servers

A privacy protection scenario for three email users ($Alice$, Bob, and $Charlie$) in a mail server G to enforce privacy protection policies under a specific purpose from different organization domain is shown as follows:

G company is a well-known mail server portal that provides email sending, receiving, and storing management services for its registered users. In order to apply for an email account from this portal, each new user has to explicitly fill in his own office profile information to this portal, including name, office phone number, office address, and working organization, etc. Furthermore, for the purposes of providing the user's personal email search and retrieval or for the management of a mail server's own business services, dynamically generated users' digitally traced information will be online extracted, (un)disclosed, and even archived in this portal during email sending and receiving activities.

The possible online digitally traced information extracted, (not-)disclosed, and archived from this mail server portal are IP address for each time the user signs in, sender's/receivers' email address(es) for each incoming/outgoing email, the titles and contents for each thread of all associated emails, etc. Of course, the mail server G does provide opt-in and opt-out mechanisms for the user to decide whether his public (or private) profile and digitally traced information can be (not-)disclosed under certain circumstances for some roles to achieve a specific purpose.

Please propose $DL + log$-based privacy protection policies that can explicitly specify ontologies and rules to satisfy the *weak DL-safeness* conditions to have the semantic enforcement of privacy protection objective via the combination of ontologies and rules [23]. The knowledge bases of ontologies and rules for two use case scenarios can be shown as: [2].

A 5-tuple term $(user(s), type(s), purpose(s), right(s), condition(s))$ is a fact shown as the P3P XML-based representation from data owner specified options on the data usage's for data user(s), where user(s) $\in$ data user ontology, type(s) $\in$ data type ontology, purpose(s) $\in$ purpose ontology; right(s) $\in$ (read,write,display,disclose,..), and condition(s) $\in$ (date,time,counter,..). Once this 5-tuple term was collected from data owner, it will be extracted and decomposed as several legal predicates that fitted into the grounding facts for the ontologies module and the

[2] In the following rules and facts, each term shown as capital letters comes from ontologies while each term shown as little letters is defined as Datalog predicates. This is the feature of a hybrid ontologies+rules combination.

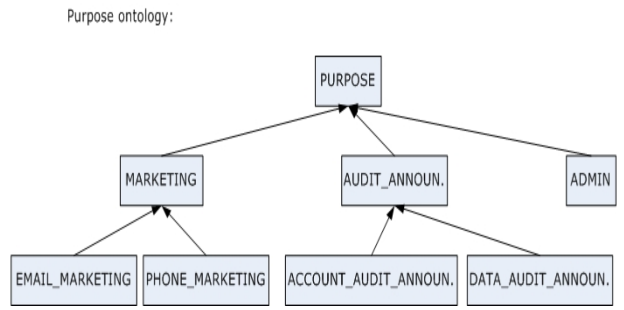


Figure 3. A purpose ontology for the classification of different data usage purposes

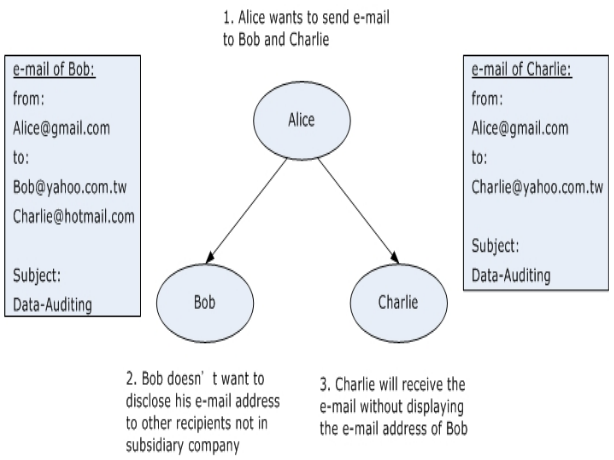
rules module to semantically enforce the privacy protection policies with respect to each data user's request.

- **Use case one scenario**: There are two organizations that share users' public profiles and digitally traced information from this mail server portal: one is a subsidiary department SD of this mail server and the other is a cooperative partner CP of this mail server. The privacy protection policies to enforce the information disclosure requests from the members of these two organizations will be quite different from service purposes or user roles perspective. Now a user $Alice \in SD$ is going to send a data auditing announcement email $\in DATA\_AUDIT\_ANNOUN.$ to both a user $Bob \in SD$ and a user $Charlie \in CP$. Under company SD internal regulation, anyone sends an email to a mailing list with multiple recipients, where email recipients $\in SD$ cannot disclose his/her email address to those people not $\in SD$ domain under any purposes. Therefore, the email recipient $Charlie \in CP$ cannot explicitly see the email address of the recipient $Bob \in SD$ in his receiving email address header(see Figure 4).

  Let $\Gamma = (\Lambda, \Delta)$ be the two components of knowledge representation from ontologies Λ module and rules Δ module:

  - Λ = ontology about information disclosure for this use case one scenario:

    Ontologies Module's Axioms:
    COMPANY $\sqsubseteq$ *PRIVATE*
    PRIVATE $\sqsubseteq$ *ORGANIZATION*
    OWNER $\sqsubseteq$ *PERSON*
    COMPANY $\overset{domain}{\longleftarrow} HAS\_COOPERATIVE \overset{range}{\longrightarrow}$ *COMPANY*
    COMPANY $\overset{domain}{\longleftarrow} HAS\_SUBSIDIARY \overset{range}{\longrightarrow}$ *COMPANY*
    HAS_COOPERATIVE $\equiv$ *HAS_COOPERATIVE*$^{-}$
    PERSON $\overset{domain}{\longleftarrow} IS\_STAFF\_OF \overset{range}{\longrightarrow}$


1. Alice wants to send e-mail to Bob and Charlie

e-mail of Bob:
from:
Alice@gmail.com
to:
Bob@yahoo.com.tw
Charlie@hotmail.com
Subject:
Data-Auditing

e-mail of Charlie:
from:
Alice@gmail.com
to:
Charlie@yahoo.com.tw
Subject:
Data-Auditing

2. Bob doesn' t want to disclose his e-mail address to other recipients not in subsidiary company

3. Charlie will receive the e-mail without displaying the e-mail address of Bob

Figure 4. A recipient B's email address cannot be disclosed to $C \in CP$ under all data usage purposes

ORGANIZATION
MAIL_TRACE $\overset{domain}{\longleftarrow} HAS\_MAIL\_TRACE \overset{range}{\longrightarrow}$
EMAIL
EMAIL $\sqsubseteq \exists$ *HAS_MAIL_TRACE_ONLINE*$^{-}$ *.O_EMAIL_SENDER*
EMAIL $\sqsubseteq \forall$ *HAS_MAIL_TRACE_ONLINE.O_EMAIL_RECEIVER*
DATA_AUDIT_ANNOUN. $\sqsubseteq$ *AUDIT_ANNOUN.*

Ontologies Module's Facts:
ORGANIZATION(G)
HAS_SUBSIDIARY(G, J-Corp.)
HAS_COOPERATIVE(G, Q-Corp.)
IS_STAFF_OF(Alice, J-Corp.)
IS_STAFF_OF(Bob, J-Corp.)
IS_STAFF_OF(Charlie, Q-Corp.)
HAS_EMAIL_ADDRESS(Alice,Alice@gmail.com)
HAS_EMAIL_ADDRESS(Bob,Bob@yahoo.com.tw)
HAS_EMAIL_ADDRESS(Charlie,Charlie@hotmail.com)
O_EMAIL_SENDER(Alice@gmail.com),
O_EMAIL_RECEIVER(Bob@yahoo.com.tw)
O_EMAIL_RECEIVER(Charlie@hotmail.com)
HAS_MAIL_TRACE_ONLINE(Alice@gmail.com, Bob@yahoo.com.tw)
HAS_MAIL_TRACE_ONLINE(Alice@gmail.com, Charlie@hotmail.com)

  - Δ = Rules about information disclosure for this use case one scenario:

    Rules Module's Rules:
    cando(?c,?b-email, display) $\Longleftarrow$
    opt-in(?b,?b-email,?p)), data-user(?c),
    data-owner(?b),
    HAS_EMAIL_ADDRESS(?b,?b-email). $\leftarrow$ *(a1)*

cando(?c,?b-email, nill) ⟸
opt-out(?b,?b-email,?p)), data-user(?c),
data-owner(?b),
HAS_EMAIL_ADDRESS(?b, ?b-email). ← *(a2)*

opt-in(?b,?b-email,?p) ⟸
IS_STAFF_OF(?b,?c1), IS_STAFF_OF(?c, ?c2),
HAS_SUBSIDIARY(?c1,?c2),
HAS_MAIL_TRACE_ONLINE(?a-email,?c-email),
O_EMAIL_SENDER(?a-email),
O_EMAIL_RECEIVER(?c-email),
data-owner(?b), data-user(?c), purpose(?p),
data-type(?b-email). ← *(a3)*

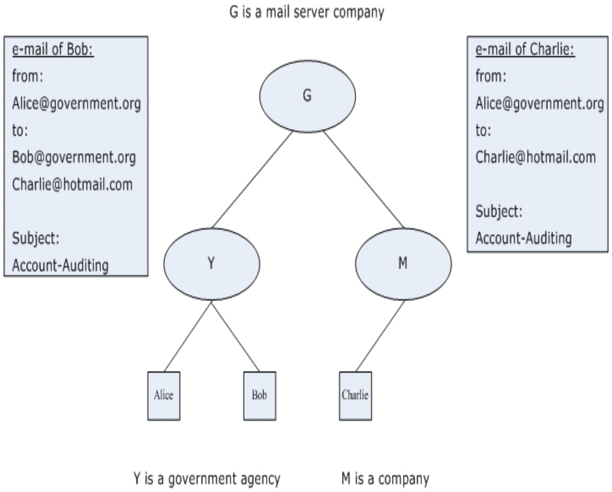
opt-out(?b,?b-email,?p) ⟸
IS_STAFF_OF(?b,?c1), IS_STAFF_OF(?c, ?c2),
HAS_COOPERATIVE(?c1,?c2),
HAS_MAIL_TRACE_ONLINE(?a-email,?c-email),
O_EMAIL_SENDER(?a-email),
O_EMAIL_RECEIVER(?c-email),
data-owner(?b), data-user(?c), purpose(?p),
data-type(?b-email). ← *(a4)*

Rules Module's Facts:
data-user(Bob), data-owner(Bob),
data-user(Charlie), data-owner(Charlie),
purpose(data-auditing),
data-type(Bob@yahoo.com.tw),
data-type(Charlie@hotmail.com),
opt-in(c,Charlie@yahoo.com,data-auditing),
cando(Bob,Charlie@yahoo.com,display),
cando(Charlie,Bob@yahoo.com.tw,nill),
opt-out(b,Bob@yahoo.com.tw,data-auditing)

From *Bob*'s side, a mail server G will be grounding rule (a4) first and then it will derive *opt-out(b,Bob@yahoo.com.tw,data-auditing)* as a conclusion. The *opt-out(..)* will become one of the facts in rule (a2) conditions once *Charlie* activates his email receiving action from mail server G to read this particular email from $Alice@gmail.com$. The recipient email address *Bob@yahoo.com.tw* will not be displayed due to the conclusion of *cando(Charlie,Bob@yahoo.com.tw,nill)* from rule (a2) due to the $nill$ access right.

From *Charlie*'s side, a G mail server does not have the constraints from *Charlie* to enforce associated privacy protection policies so *Bob* is aware $Charlie$ as one of the mailing list recipients with *Charlie@hotmail.com* in his receiving email message (see Figure 4). In the rule $(a3)$, it satisfies *weak DL-safeness* but it does not satisfy *DL-safeness* conditions because some of the variables $c1$ and $c2$ in *IS_STAFF_OF* DL predicate did not occur in any Datalog predicates.

- **Use case two scenario**: The auditing officer Bob serves in one of government auditing agencies Internal Revenue Service (IRS), where $IRS \in GOV\_AGENCY \sqsubseteq PUBLIC$. *Bob* is going to enforce a routine auditing check to a company $M \in COMPANY \sqsubseteq PRIVATE$ through its representative $Charlie$. An auditing announcement officer $Alice$ from IRS is going to send an email to a representative employee $Charlie \in M$ and other company representatives to notify the account-auditing schedule. Under government's auditing regulations, the real acting auditor Bob as one of the mailing list recipients served in IRS cannot disclose his email address in this account-auditing notification email. Therefore, a chief privacy officer (CPO) $\in IRS$ has to opt-out the acting auditor recipient Bob's email address to comply the regulations while $Alice$ is sending an account-auditing notification message (see Figure 5).



e-mail of Bob:
from:
Alice@government.org
to:
Bob@government.org
Charlie@hotmail.com

Subject:
Account-Auditing

e-mail of Charlie:
from:
Alice@government.org
to:
Charlie@hotmail.com

Subject:
Account-Auditing

Figure 5. A recipient Bob's email address $Bob@government.gov$ cannot be disclosed to $Charlie$ under auditing regulations for the purpose of delivering auditing notification email to $Charlie$

The ontologies module and the rules module for this use case two scenario are very similar to those specified in the use case one scenario except conditions for rule $(a3)$ and rule $(a4)$ are not shown as binary ontology predicates $HAS\_SUBSIDIARY(..)$ and $HAS\_SUBSIDIARY(..)$ instead they are replaced as unary ontology predicates $IRS(?c1)$ and $IRS(?c2)$ to ascertain the data owner b will *opt-in(..)*

his email address to the data user c who also serves in IRS. Otherwise, the data owner b will *opt-out(..)* his email address to the data user c who is not an IRS employee.

5 Discussion

5.1 Which Ontologies+Rules Combination?

A variety of ontologies and rules (ontologies+rules) combinations had been proposed for the past few years, such as *DLP, SWRL, AL-log, DL-log, DL+log*, and *MKNF*, etc [11] [12] [9] [23] [24] [19]. We subjectively choose *DL+log* as the ontologies+rules combination for two use case scenarios of privacy protection because *DL+log* constitutes the most powerful decidable combination of Description Logic (DL) ontologies and disjunctive Datalog rules with a *weak DL-safeness* rule condition [23].

This condition of *DL-safeness* can be expressed as follows: every variable occurring in an atom with a DL predicate must occur in an atom with a Datalog predicate in the body of the rule [24]. In other words, the *DL-safeness* condition ensures that each rule variable must occur in one of the Datalog predicates. In *DL+log* ontologies+rules combination, *DL-safeness* can be weakened as *weak DL-safeness* without losing its nice decidable computational properties where a Datalog rule with only on the head variables of the rule imposed *DL-safeness* condition [23], e.g., every head variable of Datalog rule must appear in at least one of the atoms in a Datalog predicate.

A Semantic Web Rule Language (SWRL) used to be a semantic web language for the combination of ontologies+rules [12]. But the complexity of reasoning for query of SWRL-based ontologies+rules is undecidable, which prevents people from using this combination without hesitation. Decidability of reasoning is a crucial issue in systems when we combine DL-based knowledge bases (KBs) and Datalog rules together [20] [22]. The loose integration between DL-based KB and rule-based KB with *weak DL-safeness* conditions apply to all of the rules in a privacy protection policy set that guarantee the semantic enforcement of privacy protection policies to be a decidable decision process.

5.2 Privacy Protection Language and Policy

A privacy protection language and policy was proposed by Karjoth, G. as an extending Flexible Authorization Framework (FAF) with grantors and obligations [15]. In this extended FAF approach, a privacy control language includes user consent, obligations, and distributed administration. There are several issues they did not exploit in their approach, shown as follows:

1. They did not explicitly separate the ontologies module and rules module in their policy specification so the rules to enforce privacy protection policies have to be classified as four categories: direct authorization rules, derived authorization rules, decision rules, and integrity rules. In our approach, the decisions for the derived authorization rule can be enforced directly from the reasoning of ontologies using class and subclass subsumption relationships. Then the rules module only has to deal with the final permission of information disclosure.

2. They did not really demonstrate how to achieve the authorization decision of private information disclosure using the combination of hierarchy of groups, data objects, and purposes. We explicitly show this authorization decision can be obtained by the ontology merging techniques from our three ontologies, e.g., data subject ontology, object ontology, and purpose ontology.

3. They did not model and enforce the disclosure of data between enterprises, i.e., exporting and importing data with their associated privacy policy from/into a system. We are aware that this problem can be solved by using ontologies merging and rule composition techniques across multiple domains [5].

4. Finally, they did not consider the profile information disclosure as well as the digitally traced information disclosure and this will be an emerging research area for Web 2.0 and Web 3.0 privacy protection. In our above mail server use case, we demonstrated how the personal profile information disclosure opt-in/opt-out selection influences the later on digitally traced information disclosure.

6. Conclusion and Future Prospects

There are several challenges for us to elaborate the semantic web core technologies on modeling of privacy protection's policy representation and enforcement. At this moment, we are not quite sure which ontologies+rules combination will be the most appropriate one under certain information usage purposes and conditions [9] [19] [23] [24].

In summary, we express and enforce all profile information and digital traces with associated disclosure policies using a specific ontologies+rules combination on Web 3.0, e.g., $DL + log$. This information modeling structure and access mechanism will be quite different from Web 1.0 and Web 2.0, where profile information will be defined as relational database tables in the deep web, and digital traces for recording each user's surfing activities will be defined and collected as an unstructured weblog. On the Web 3.0 information cyberspace, we might face all personal profile information as well as associated digital traces are modeled as a ontologies+rules combination with semantic query as the only feasible access mechanism; then the challenge for semantic representation and enforcement of privacy protection policies just begins.

References

[1] A. H. Anderson. A comparison of two privacy policy languages: Epal and xacml. In *Proceedings of the 3rd ACM Workshop on Secure Web Services (SWS'06)*, pages 53–60. ACM, 2006.

[2] I. A. Antón et al. A roadmap for comprehensive online for privacy policy management. *Comm. of the ACM*, 50(7):109–116, July 2007.

[3] G. Antoniou et al. Rule-based policy specification. In T. Yu and S. Jajodia, editors, *Secure Data Management in Decentralized Systems*, pages 169–216. Springer, 2007.

[4] M. Blaze, J. Figenebaum, and M. Strauss. Compliance checking in the policymaker trust management system. In *Proc. of the Financial Cryptography*, LNCS 1465, pages 254–274. Springer, 1998.

[5] A. P. Bonatti, S. D. C. di Vimercati, and P. Smarati. An algebra for composing access control policies. *ACM Trans. on Information and Systems Security*, 5(1):1–35, February 2002.

[6] A. P. Bonatti et al. Semantic web policies - a discussion of requirements and research issues. In *3rd European Semantic Web Conference (ESWC 2006)*, Budva, Montenergro, June 2006.

[7] L. Cranor et al. The platform for privacy preferences (p3p) 1.0 (p3p 1.0) specification, 2002. http://www.w3.org/P3P/.

[8] L. Cranor, M. Langheinrich, and M. Marchiori. A p3p preference exchange language 1.0 (appel 1.0), 2002. http://www.w3.org/TR/P3P-preferences/.

[9] M. F. Donini et al. *AL*-log: Integrating datalog and description logics. *Journal of Intelligent Information Systems*, 10(3):227–252, 1998.

[10] S. Fischer-Hübner. *IT-Security and Privacy - Design and Use of Privacy-Enhancing Security Mechanisms*. LNCS 1958. Springer, 2001.

[11] N. B. Grosof et al. Description logic programs: Combining logic programs with description logic. In *World Wide Web 2003*, pages 48–65, Budapest, Hungary, 2003.

[12] I. Horrocks et al. Swrl: A semantic web rule language combing owl and ruleml, 2004. http://www.w3.org/Submission/SWRL/.

[13] S. Jajodia et al. Flexible support for multiple access control policies. *ACM Trans. on Database Systems*, 26(2):214–260, June 2001.

[14] L. Kagal, T. Finin, and A. Joshi. A policy based approach to security for the semantic web. In *International Semantic Web Conference (ISWC) 2003*, LNCS 2870, pages 402–418, 2003.

[15] G. Karjoth and M. Schunter. A privacy policy model for enterprises. In *15th IEEE Computer Security Foundations Workshop (CSFW)*. IEEE, June 2002.

[16] G. Karjoth, M. Schunter, and M. Waidner. Platform for enterprise privacy practices: Privacy-enabled management of customer data. In *2nd Workshop on Privacy Enhancing Technologies (PET)*, LNCS. Springer, 2002.

[17] N. Li, T. Yu, and A. I. Antón. A semantics-approach to privacy languages. *Computer Systems and Engineering (CSSE)*, 21(5), Sep. 2006.

[18] J. Maluszynski. Hybrid integration of rules and dl-based ontologies. In J. Maluszynski, editor, *Combining Rules and Ontologies. A survey*, pages 55–72. EU FP6 Network of Excellence (NoE), Feb. 2005. REWERSE.

[19] B. Motik et al. Can owl and logic programming live together happily ever after? In *5th International Semantic Web Conference (ISWC) 2006*, LNCS 4273, Athens, GA, USA, Nov. 2006.

[20] B. Motik, U. Sattler, and R. Studer. Query answering for owl-dl with rules. In *3rd International Semantic Web Conference (ISWC) 2004*, LNCS 3298, pages 549–563. Springer, 2004.

[21] J. Park and R. T. Sandhu. The $ucon_{ABC}$ usage control model. *ACM Trans. on Information and System Security*, 7(1):128–174, 2004.

[22] R. Rosati. On the decidability and complexity of integrating ontologies and rules. *Web Semantics: Science, Services and Agents on the World Wide Web 3*, pages 61–73, 2005.

[23] R. Rosati. *DL*+log: Tight integration of description logics and disjunctive datalog. In *Proc. of the 10th International Conference on Principles of Knowledge Representation and Reasoning (KR)*, 2006.

[24] R. Rosati. Integrating ontologies and rules: Semantic and computational issues. In *Reasoning Web 2006*, LNCS 4126, pages 128–151, 2006.

[25] G. Tonti et al. Semantic web languages for policy representation and reasoning: A comparison of kaos, rei, and ponder. In *2nd International Semantic Web Conference (ISWC) 2003*, LNCS 2870, pages 419–437, 2003.

[26] S. D. C. d. Vimercati et al. Access control policies and languages in open environments. In T. Yu and S. Jajodia, editors, *Secure Data Management in Decentralized Systems*, pages 21–58. Springer, 2007.

[27] D. J. Weitzner et al. Creating a policy-aware web: Discretionary, rule-based access for the world wide web. In E. Ferrari and B. Thuraisingham, editors, *Web and Information Security*, pages 1–31. Idea Group Inc., 2006.

[28] Y. C. T. Woo and S. S. Lam. Authorization in distributed systems: a new approach. *Journal of Computer Security*, 2(2-3):107–136, 1993.

[29] T. Yu, A. N. Li, and I. Antón. A formal semantics for p3p. In *ACM Workshop on Secure Web Services*, Fairfax, VA, USA, Oct. 2004. http://citeseer.ist.psu.edu/750176.html.