Representation and Inference of Privacy Risks Using Semantic Web Technologies

Renata Vieira
PUCRS
Ipiranga Av. 6681, FACIN
CEP 90619-900
Porto Alegre, Brazil
renata.vieira@pucrs.br

Douglas da Silva
PUCRS
Ipiranga Av. 6681, FACIN
CEP 90619-900
Porto Alegre, Brazil
douglas.silva@cpph.pucrs.br

Tomas Sander
Systems Security HP Lab
5 Vaughn Dr,
Princeton, NJ 08540, USA
tomas.sander@hp.com

ABSTRACT
This poster discusses domain ontologies on the privacy field for automatic risk identification and project qualification. It presents an ontology model for describing risks as an interpretation of privacy policies contextualized in project specifications.

Categories and Subject Descriptors
K.4.1 [**Computers and Society**]: Public Policy Issues – *privacy.*

General Terms
Management, Reliability, Security, Legal Aspects.

Keywords
Ontologies, privacy, accountability, risk assessment.

1. INTRODUCTION
Attention to privacy legislation is an important issue in IT (information technology) projects to avoid lawsuits and loss of consumer trust [1]. The privacy domain in IT and accountable privacy management (APM) in organizations are explored in several works, as follows. In [2] a taxonomy describes concepts such as collection, processing, dissemination and invasion of information. Knutson [3] presents basic principles to create privacy awareness in software projects. This is done through the identification of privacy goals that fulfill legal obligations, the definition of a privacy core team with technical and legal experts and the creation of guidelines to help developers to become independent from the privacy experts. Similar concerns for software design are endorsed within other work on privacy awareness [4][5]. The KAoS Policy Ontologies (KPO) [6] defines concepts such as actions, actors, groups, places, entities related to actions, and policies. An integration of policies relating several aspects of security is proposed in [7], including authorization and privacy into semantic web services. The BDSG ontology, mapping law statements to a machine interpretable language is presented in [8] as a way to enforce privacy in enterprises using ontologies to generate XACML [9] policies. Hecker [10] argues that ontologies on the privacy field must enable interoperability, determine the privacy level of a transaction and guide the implementation of privacy functionalities without requiring expertise from the domain. Hecker creates an ontology using terms from privacy notions and concepts from the European Parliament Directive 95/46/EC11. Hu [11] proposes that the semantic model for EPAL privacy policies can be expressed as a variety of combinations of ontologies and rules.

Considering the importance of a proper representation of relevant rules formulated for handling personal information, described in laws, guidelines, policies, and other normative sources, we propose an ontology describing data privacy risk related concepts, based on project actions. In particular, we address the appropriate usage of information under a set of rules based on the identification of privacy risks when dealing with sensitive information. An OWL ontology infers risks in terms of actions of software projects and the effect of these actions over sensitive information.

2. REPRESENTING PRIVACY RISKS
Our model includes concepts, such as personally identifiable information (PII), sensitive information, user actions, location, and risk levels. Restrictions and constructors to classify actions in risk actions are described through object properties. Figure 1 shows some of the relevant identified classes.

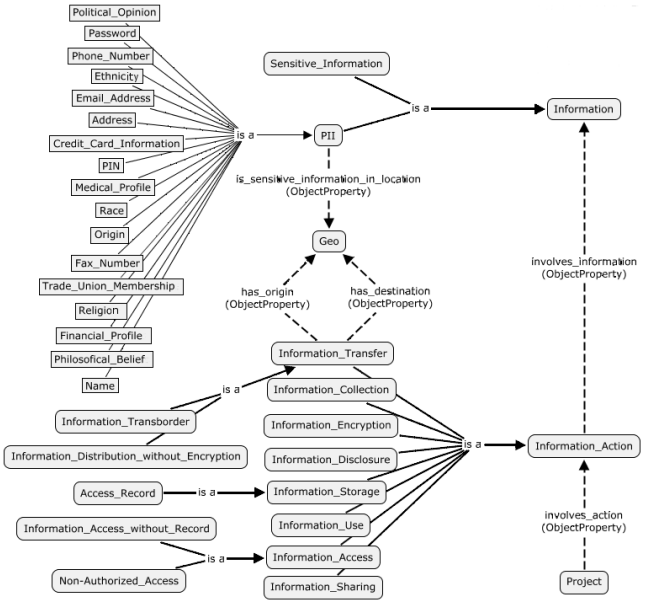

Figure 1. Privacy domain classes.

These are common concepts in documents, laws and guidelines of the domain for privacy assurance and accountable software

development[1]. Properties are used to relate actions to contextual information, examples are:

```
Information Transfer has origin Country
Information Transfer has destination Country
PII is sensitive information in location Geo
Action involves information Information
Action has secondary action Action
```

Based on Argentina's provision [12], which presents a classification of risks as low, moderate and high, we exemplify risk inference as follows. Figure 2 shows Risk_Action classes.

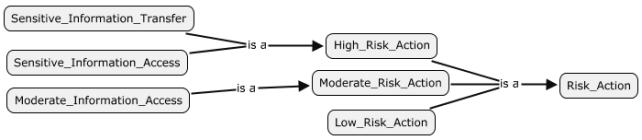

Figure 2. Risk classes.

In our model, privacy risks are inferred from project related information. Consider an instance of a project in the ontology, which refers to actions that manipulate information, such as name, address and social security number. This project instance is asserted using the relations: *involves action, involves information, has origin, has destination,* and *is sensitive information in location,* as shown below.

```
project involves action
          information access
          information transfer
information access involves information
          name
          address
          social security number
information transfer involves information
          name
          address
          social security number
information transfer has origin
          USA
information transfer has destination
          Portugal
social security number is sensitive information in location
          USA
```

With these facts asserted in the ontology, an instance involving social security number will be classified as Sensitive Information in the USA as a consequence of a SWRL codified rule. In this way, the action will be inferred as high risk action: Sensitive Information Access and Sensitive Information Transfer.

3. FINAL REMARKS

Ontologies on the privacy domain are useful to provide ways to share vocabulary and better understand a particular domain and its related concepts. Our research is aimed at building models that infer risks automatically from the specification of project features. Such knowledge intensive areas require advanced knowledge management technologies. A privacy core team is necessary to create and maintain such systems based on dynamically changing knowledge. Our model presents concepts and relations where actions involving data related to personal information and their contexts (time, place) are related to risks. The proposed approach is intended to guide managers with risk assessment. The model is also designed to help privacy experts to formalize risky situations in organizations. As future work we plan to map our concepts to other similar ontologies, linking for instance, our action concepts to action concepts of KAoS. We are also considering the processing of textual knowledge sources such as laws and guidelines to ease the identification of relevant domain information.

[1] A hyperbolic view of the ontology is available at http://www.inf.pucrs.br/~ontolp/Visualizacao/Privacy_Risks/Privacy_risks.html.

6. ACKNOWLEDGMENTS

This paper was done in cooperation with Hewlett-Packard Brasil Ltda. using incentives of Brazilian Informatics Law (Law nº 8.2.48 of 1991).

7. ADDITIONAL AUTHORS

Additional authors: Alexandre Agustini (PUCRS), Caio Northfleet (HP), Fernando Castilho (PUCRS), Mírian Bruckschen (PUCRS), Patrícia Pizzinato (PUCRS), Paulo Bridi (PUCRS), Prasad Rao (HP), Roger Granada (PUCRS).

8. REFERENCES

[1] Mont, M., Thyne, R.: Privacy policy enforcement in enterprises with identity management solutions. In: PST '06, vol. 380, pp. 1--12. ACM, New York (2006).

[2] Solove, D. J.: A Taxonomy of Privacy. University of Pennsylvania Law Review, vol. 154, no. 3, p. 477, (2006).

[3] Knutson, T. R. 2007. Building Privacy into Software Products and Services. IEEE Security and Privacy, vol. 5, no. 3, pp. 72--74 (2007)

[4] Duncan, G.: ENGINEERING: Privacy By Design. Science 317 (5842), 1178, (2007)

[5] Ye, X., Zhu, Z., Peng, Y., Xie, F.: Privacy Aware Engineering: A Case Study. Journal of Software, vol. 4, no. 3, pp. 218--225 (2009)

[6] Bradshaw, J. et al.: Representation and reasoning for DAML-based policy and domain services in KAoS and nomads. In: AAMAS '03. Melbourne, Australia (2003)

[7] Kagal, L., Paoucci, M., Srinivasan, N., Denker, G., Finin, T., and Sycara, T.: Authorization and Privacy for Semantic Web Services, In: AAAI Spring Symposium on Semantic Web Services (2004)

[8] Abou-Tair, D.D., Berlik, S., Kelter, U.: Enforcing Privacy by Means of an Ontology Driven XACML Framework. In: IAS 2007, Third International Symposium on Information Assurance and Security, pp. 279--284 (2007)

[9] OASIS XACML Technical Committee.: eXtensible Access Control Markup Language (2003)

[10] Hecker, M., Dillon, T. S., Chang, E. Privacy Ontology Support for E-Commerce, IEEE Internet Computing, vol. 12, no. 2, pp. 54--61 (2008)

[11] Hu, Y., Guo, H., and Lin, A. G.: Semantic Enforcement of Privacy Protection Policies via the Combination of Ontologies and Rules. In: SUTC 2008, vol. 00. IEEE Computer Society, Washington, DC, pp. 400--407 (2008)

[12] Ministerio de Justicia, Seguridad e Derechos Humanos, Presidencia de la Nación Argentina. Dirección Nacional de Protección de Datos Personales (2006)