Position Paper: Privacy Policies as a Component of Policy-enabled Governance

M. Hondo, T. Nadalin, R. Nagaratnam (IBM Software Group)
M. Kudoh, G. Karjoth, B. Pfitzmann, M. Schunter (IBM Research)
9/12/2006

1 Introduction

Publishing a privacy policy and enforcing the stated privacy obligations are an important Corporate Governance requirement for industry as well as government organizations. Effective privacy enforcement and compliance often includes several aspects of access control that are not covered by today's role-based access control standards. Examples of this include a notion of purpose, or the ability to compare different policies. There is also the need to address access issues in user centric systems that store personal data. In recent years, substantial research has been done in the area of privacy enforcement but we are not left with a good answer for overall governance.

In this position paper, we outline requirements for languages addressing privacy protection and show why a privacy-only effort is too short sighted. Instead, a language should aim at formalizing broader governance requirements. We propose a roadmap for how to move towards a language that includes privacy enforcement as one component of a broader data governance perspective.

2 Requirements for Privacy Languages

Each existing privacy language can formalize some of the requirements associated with the handling of protected data. None covers the complete set of requirements for privacy languages or the broader governance requirements. In the sequel, we sketch out the detailed requirements that a language should satisfy.

2.1 Expressive and Unambiguous

The main requirement for any privacy language is that the language must be able to express access constraints on who is allowed to perform which actions on which resources for what purpose as well as to provide a mapping (deployment) into enforcement systems.

A language that would be capable of providing privacy protection within an overall corporate governance effort needs to be expressive enough to formalize the complex handling requirements in today's business. Elements for expressivity would include access rules based on pre-conditions for instance-based decisions and obligations to mandate actions before or after an access. Furthermore, it should be applicable to the following areas of data handling:

- Processes (after access, I should send a notifications)
- Systems (what systems can data be stored on)
- Processing (How should data be handled? Is encryption in transit needed? What are trusted entities?)
- Access control (access control)
- Access and audit (what audit possibilities exist)

In order to be clear and unambiguous, a language needs to have well-defined semantics. This means that given a specified way to handle data in the scope of the policy, the policy should unambiguously define the permitted handling of this data.

2.2 Composable and Comparable

It is a fundamental assumption that any privacy language is composable and comparable. Composable means that a policy can be composed of sub-policies while being able to resolve conflicts. Comparable means that enterprises can tell whether one policy is at least as restrictive as another policy with respect to given data.

An essential element to enable composition and comparability is policy scoping by means of ontologies. This means that each policy defines what types of handling and data are covered by the given policy. Another element is a policy algebra where policies can be combined using logical connectors (AND, OR, NOT ... as well as XACML's `First-Applicable' operator) such that the result is again a policy.

2.3 Scope: Corporate Governance

We believe that a ***privacy-only*** language will not be broadly implemented. Enterprises have a variety of data handling requirements. It is unlikely that enterprises will deploy separate infrastructures for each aspect. As a consequence, a policy language should also cover the context of corporate governance while clearly defining its semantics for handling information with privacy constraints.

2.4 Parts of the Language/Framework

We believe that a suitable language comes in three parts that are largely independent (but composable):

- A *core language* for the policy that includes abstract concepts such as ”purpose” and ”obligation” and provides a method for representing abstract and concrete privacy expressions.
- *Binding mechanisms* that associate the core language with other entities such as messages or services; these are partially dependent on underlying messaging or data access mechanisms (e.g., there will be one for web services).
- *Ontologies or vocabularies*. These consist of sector-independent base ontologies and sector-specific refinements. Sector-independent parts contain, e.g., definitions of data deletion and notification requirements. Common purposes such as ”marketing” or ”improving our systems” fall in the middle, while, e.g., concrete data categories or people classifications might be specific to sectors such as medical, financial, etc.

2.5 Compatible with Existing Standards

In order to facilitate broad adoption, the language needs to be compatible with existing standards. We see the following set of standards where compatibility is essential:

- HTTP – transmission of web-pages.
- SOAP – the messaging standard in a web-services environment
- WS-Policy family for assertions, semantics and attachment mechanisms
- XACML for the core language (if augmented by suitable profiles) and an underlying framework for terminology (such as “PEP” and “PDP”)

- P3P – Policies need to be transformable from P3P into the new language, as well as, with suitable abstractions, from the new language to P3P. This concerns both the core language and ontologies.

Remark: It is still an open question whether a privacy policy language shall support `upward inheritance'; see the work of Barth and Mitchell for example.

3 Standardization Roadmap

We believe that the overall goal of a language for defining data handling requirements needs to be decomposed into several horizons that should build on each other in order to provide building blocks in the evolution of policy-enabled corporate governance.

3.1 Engaging People - User-centric Policy Negotiation

Foundation of the standardization roadmap is a consented governance policy. Security practitioners know, people adapt and they are creative, that's how the species survives. Any long term privacy solution must engage not alienate people. Key to simplification will be consumability and utility. The creation of a consumable governance policy requires legal analysis and end user interaction. By means of notice, negotiation, and consent, individual end-user should be able to tailor the policy to fit their personal privacy and service preferences. The goal of the consecutive standards is then to enforce and propagate such a consented policy. A standard in this space is P3P. Other ways to gather consent are privacy notices and collection of opt-in and opt-out choices.

3.2 Privacy-enhanced Access Control

The first step is to define a core privacy-enhanced access control language. The first sub-step is to identify the set of requirements to be covered. Since XACML is the only broadly supported standard in this space, the next sub-step would be to examine how to satisfy the given requirements with an extension of XACML. A first step in this direction is a privacy profile of XACML, which goes far beyond what is currently defined by OASIS.

3.3 Access and Privacy Control for Web-services

The second step builds on the first and enables enterprise-internal use of privacy enhanced access control for web-services. A standardized binding mechanism is needed that defines how policy-protected data can be transmitted using SOAP. This standard may leverage WS-PolicyAttachment.

3.4 Interoperability and Deployment by means of Ontologies

XACML does not have ontologies or other scoping tools. As a consequence, it is unclear what is covered by an XACML policy. The next step would be to define a scoping/ontology format for XACML and standardize how XACML policies can refer to the corresponding ontology.

3.5 Policy Exchange, Negotiations and Consent

The next step is to allow for transmission and negotiation of policies. By using ontologies, two enterprises can agree on common terminology. By attaching policies to web services artifacts, policy negotiation is enabled, at least in the web services case. By using policy comparison, for instance an enterprise sending data can determine whether the receiving enterprise protects the data at least as good as the sending enterprise.

User interaction may be necessary in order to, for instance, obtain consent for the release of some piece of that user's identity information or to collect a piece of identity information requested by another provider.

3.6 Governance beyond Access

The final step is to examine how other aspects of data handling can be formalized. This will enable enterprises to express other data handling requirements that go beyond access restrictions.

4 Conclusion: Making a governance solution happen

Currently, it is difficult for businesses to assess how a privacy-only language will help address the privacy governance problems they face. Their governance requirements include not only the annotation of information as “needing privacy” but also include enforcement of policies to protect information and provide users with appropriate access to the information they require. It is not sufficient to provide privacy by locking down all access. The challenges lie in creating systems in which privacy requirements are considered and integrated into the decisions made by people and systems every day. Understanding privacy issues will help us create better applications and systems for general use and this is where the cost can be justified and recovered.

As citizens, none of us want to be restricted from access to public information because the systems can’t discriminate at the proper level of granularity. The challenge before us is to design FOR privacy-enablement, not to design new systems that only do privacy, but the former is a more difficult task than the latter, and so tends to not be undertaken.

We propose a straw man how to make our roadmap happen.

Starting point: Use and extend existing access control to meet basic privacy requirements identifying new requirements for privacy management in the context of an overall governance language. Make these privacy requirements implementable: Simple privacy notices translate to simple policies enforceable by access control. Focus on making a core set deployable in many products/platforms thus enforcing the same policy subset on many platforms. Integrate privacy protection into existing access control systems (language as well as enforcement).

Mid point: Provide businesses and Industry consortia with a map of product support including the commitment of vendors and the results of demonstrated interoperability events. Continue with a cross Industry push to constrain niche standard development. Try to limit overly academic efforts while focusing on efforts that embed simple, enforceable policies in other deployed industry efforts standards (examples include standards for health records or federated identity management).

End point Goals: Update the language to meet all requirements. This will be based on new research such as complex decision making involving multiple factors and considerations (tradeoffs between access time and amount of protection end to end). Understand the blockers for optimizing systems --- identify where the extra cost is, how to address the problems of consistency between access control and privacy policy, and how systems can help people protect their privacy but still be global in their thinking and actions. Evaluate where network level optimizations can provide optimal/accountable alternatives to enable more confidence in global participation in a wired world.