Data-centric Privacy Policies for Smart Grids

Sebastian Speiser and Andreas Harth
The AAAI 2012 Workshop on Semantic Cities
Toronto, Canada July 23rd, 2012
Agenda

- Motivation
- Data-centric Policies
- Policies restricting Policies
- Patterns
- Sticky Policies
- Conclusions
MOTIVATION
Motivation – Example

Alice Energy Consumption Data → access → storing → deidentification → Deidentified Energy Consumption Data
Motivation – Example

Subject to privacy policy

Alice Energy Consumption Data → access → Deidentified Energy Consumption Data

Access Control: Only her energy provider may access the data
Motivation – Example

Subject to privacy policy

Obligation:
Stored data must be deleted after one year
Motivation – Example

Subject to privacy policy

Policy restrictions: Deidentified data may also be shared under a policy that allows only non-commercial usage.
Motivation

- Privacy policies can restrict:
  - Access to data
  - Usages of data, including specification of obligations
  - Policies for publishing or sharing derived data

- Formal policies help to automate compliance checks

- Challenge: large number of heterogeneous entities
  - No central view or control of processes
  - Individual privacy requirements differ
  - No central data storage
  - Intensified when Smart Grid is coupled with other Smart City systems
DATA-CENTRIC POLICIES
## How to express usage policies?

### Process-centric view
- **On system level**
  - Sharing of consumption records must be approved by their owners
- **On process level**
  - [a process that is a usage for non-commercial purpose or a sharing with someone, who employs] is allowed
- **In a central store**

### Data-centric view
- **On instance level**
  - Alice specified that her consumption record may be shared with Bob
- **On action level**
  - Usage for non-commercial purposes and sharing with same policy are allowed
- **Attached to artefact**
Advantages of Data-centric Policies

- Process in which artefact is used can be partially unknown
  - Useful if service is provided by network of providers
  - Adaptive to process changes (innovation)

- Intermediate artefacts have explicit policies
  - Policy can be passed with artefact to third party

- Each artefact can have its own policy
  - Fine granular usage restrictions dependent on data owner in contrast to a law applicable to all
Data-centric Policies

- Policies describe sets of compliant usages, i.e., restrictions on the actions and the policies of generated artefacts.
- Actions using an artefact must comply to artefact’s policy.
- Local view enabled by two assumptions:
  - used artefacts have correct policies;
  - generated artefacts are used in compliant way.
Usage Model for Data-centric Policies

- Based on Open Provenance Model (OPM)
  - **Usage**: using an artefact for a given purpose
  - **Derivation**: generate new artefacts that again have a policy
  - **wasTriggeredBy**: action can only start after other action started
  - **Process**: chosen as term to align OPM, but treated as atomic

![Diagram of Usage Model]

- **Policy**
  - hasPolicy
  - wasGeneratedBy

- **Artefact**
  - used
  - wasGeneratedBy

- **Derivation**

- **Process**
  - wasTriggeredBy
  - performedBy
  - performedAt

- **Actor**
  - hasPolicy

- **Time**
  - performedAt

- **Usage**
  - hasPurpose

- **Purpose**

- **class**
- **subclass of**
- **property**
  - arrow start: domain
  - arrow end: range
POLICIES RESTRICTING POLICIES
Policy Restrictions on other Policies

- Policies of artefacts generated by a derivation are dependent on policies of used artefacts
- Inheritance
  - Derived artefacts have exact same policy as inputs
- Name-based restrictions
  - Possible policies for derived artefact are listed
- Content-based restrictions
  - Possible policies for derived artefact are described
Inheritance of Policies

- Generated artefacts inherit the policies of used artefacts

- Problem: after transformation policies can become …
  - … more relaxed (e.g., after anonymisation),
  - … more strict (e.g., after combination with other data)
  - … incompatible (e.g., combining CC BY-SA and CC BY-NC-SA)

- Example: GPL
  - GPL is inherited by derived code artefacts (viral)
Name-based Policy Restrictions

- Policy of used artefact specifies exhaustive list of admissible policies for derived artefacts
- Example: Creative Commons ShareAlike licenses
  - Name-based restrictions are not intended, as they prevent compatibilities of licenses with same meaning, but different names (Lessig, Creative Commons)*

- Even more relevant for privacy policies, as we cannot assume canonical names

*: Lessig, L.: CC in Review: Lawrence Lessig on Compatibility. Available at http://creativecommons.org/weblog/entry/5709, 2005
Content-based Policy Restrictions

- Policy of used artefact specifies restrictions on the usages allowed by policies of derived artefacts
- Restrictions: must allow at least/at most certain usages
  => conditions on containment between policies
- Restriction can refer to policy itself (self-referential)
- Examples
  - anonymising artefact enables arbitrary non-commercial usages
  - derived data must be stored so that notification of usage is required and further derivations have the same terms
  - policy of confidential artefact requires that it is used only in documents with policies as restrictive as the original policy
Policy Language

- Policies as Datalog query with one head variable, e.g.:
  \[ \text{UNC}(x) \leftarrow \text{Usage}(x) \land \text{hasPurpose}(x, p) \land \text{NonCommercial}(p) \]

- Compliant policy subjects: symbols that are query answers
- Each policy is identified by an individual (policy name)

- containedIn relation for content-based policy restrictions
  - Holds between policy names if containment holds for their queries
  - containedIn is maximised for increased compatibility
Data Sharing / Rights Delegation

- Rights holder can share the data with further parties under same or more restricted policy

- \[ P1(x) \leftarrow \text{Usage}(x) \lor (\text{Sharing}(x) \land \text{wasGenBy}(a, x) \land \text{hasPolicy}(a, p) \land \text{containedIn}(p, P1)) \]

- Variations
  - further restrictions on usage or sharing (actor, time, purpose, …)
  - containment in other, more restricted policies
  - limited depth by decreasing sharing count in each policy restriction
Anonymisation

- Very similar to rights delegation
- After anonymisation typically more usages are allowed

\[ \text{PO}(x) \leftarrow \text{Anonymisation}(x) \land \text{wasGenBy}(a, x) \land \text{hasPolicy}(a, p) \land \text{containedIn}(p, \text{PA}) \]
\[ \text{PA}(x) \leftarrow \text{Usage}(x) \]

- Variation: require minimum rights granted by policy
  - containedIn(\text{PM}, p)
  - E.g., non-commercial usage must be allowed
Obligations and Time Spans

- Obligations allow temporary policy violations
  - E.g., data must be deleted within one year
- Obligations without time restrictions are ineffective

\[
PD(x) \leftarrow \text{Storing}(x) \land \text{wasGenBy}(a, x) \land
\text{wasTriggeredBy}(d, x) \land \text{Deletion}(d) \land \text{used}(d,a) \land
\text{performedAt}(d, t) \land t \leq "2012-12-31".
\]

- Data-centric policies need absolute time restrictions
  - Otherwise: just store a new copy and the timer is reset
STICKY POLICIES
Attaching policies in an HTTP-based Architecture

- RIF identifies rules with IRIs
- We adopt the Linked Data principle: Resolving the IRI of a policy should return its definition

- For Linked Data, we have identifiers for documents assign policy in RDF using the hasPolicy property

- For other resources, we can use the HTTP Link Header
  \[\text{Link: } <http://ex.org/pols#P1>; \text{ rel=policy}\]
CONCLUSIONS
Conclusions

- Data-centric usage policies are suitable for Smart Grid, as
  - data is used in different contexts across dynamic provider networks;
  - each data artefact can have its own usage restrictions;
  - derived artefacts can be releasable under restricted policies.

- Content-based restrictions on other policies increase the
  compatibility by overcoming naming dependencies

- Stickiness of policies to transport them together with data

- Advantages also apply when Smart Grids are integrated
  with further Smart City systems
Thank you for your attention.

www.ksri.kit.edu  www.aifb.kit.edu
Data-centric Privacy Policies for Smart Grids

Sebastian Speiser and Andreas Harth
The AAAI 2012 Workshop on Semantic Cities
Toronto, Canada       July 23rd, 2012