A Privacy Preference Ontology (PPO) for Linked Data

Owen Sacco and Alexandre Passant
DERI, NUI Galway
firstname.lastname@deri.org

Tuesday, 29th March 2011
LDOW2011, Hyderabad India
Context

- **Linking Open Data community**
  - Encourages people to publish formatted data on the Web
  - The data does not include any metadata that describes privacy restrictions
  - Hence: the data is easily accessible

- **Access Control Lists (ACL)**
  - Specify access control to the whole RDF “document”
  - Described using Web Access Control (WAC) Vocabulary
    - Read / Write / Control
Protecting Data
- Does not only mean granting full access or not
- Requires fine-grained access control mechanisms

Current Linked Open Data environments:
- Lack mechanisms for creating fine-grained access control
- Discourages people and organisations to publish sensitive personal information
Use Cases

- Protecting a FOAF based Social Network where users:
  - Would feel more confident when publishing their personal information
  - Would be in full control
    - Which specific personal information can be shared
    - Who can access their data
  - Example: A user wants to restrict a phone number to whoever works at DERI
Use Cases

- Protecting sharing of microblog posts in SMOB
  - Microblogs in SMOB: described in RDF using ontologies such as FOAF and SIOC
  - SMOB provides tagging posts with concepts from GeoNames and DBpedia
  - Fine-grained privacy settings are required to restrict access to:
    - User’s specific information
    - Posts to users that have similar interest to the annotated concept
  - Example: A user wants to restrict a microblog post tagged with the concept of Linked Data to users that have a similar interest
A light weight vocabulary for defining fine-grained privacy preferences for RDF data

- The lightweight vocabulary should be able to restrict:
  1. A particular statement; or
  2. A group of statements (i.e. as an RDF graph); or
  3. A resource – either as a subject or as an object of a particular statement

- The Web Access Control (WAC) vocabulary is used to describe the access privilege to the data:
  - Read
  - Write
  - Control
A privacy preference contains:

- Which resource, statement or graph must be restricted
- A condition that must be satisfied
- The access control privilege (defined using WAC)
- A SPARQL query that tests whether a user requesting information matches a graph pattern

Example:
- Restrict a microblog post that contains a particular tag to the users who are interested in that tag.
Privacy Preference Ontology

- **ppp:PrivacyPreference**
- **ppp:hasAccessSpace**
- **ppp:AccessSpace**
- **ppp:hasAccessQuery**
- **rdfs:Literal**

- **ppp:hasAccess**
- **ppp:hasCondition**
- **ppp:appliesToResource**
- **ppp:appliesToStatement**
- **ppp:appliesToNamedGraph**

- **acl:Access**
- **ppp:Condition**
- **rdfs:Resource**
- **rdf:Statement**
- **trix:Graph**

- **ppp:resourceAsSubject**
- **ppp:resourceAsObject**
- **ppp:classAsSubject**
- **ppp:classAsObject**
- **ppp:hasLiteral**
- **ppp:hasProperty**

- **rdfs:Resource**
- **rdfs:Resource**
- **rdfs:Class**
- **rdfs:Class**
- **rdfs:Literal**
- **rdf:Property**

**Restrictions to**  **Conditions on**  **Access Test Queries**  **Access Control Privileges**

**Online:** [http://vocab.deri.ie/ppo#](http://vocab.deri.ie/ppo#)
Privacy Preference Ontology

Restrictions to:
- ppo:appliesToResource: restricts a resource using its URI
- ppo:appliesToStatement: restricts a particular triple by specifying the subject, predicate and object
- ppo:appliesToNamedGraph: restricts a group of statements which are identified with a URI
Privacy Preference Ontology

- **Conditions - ppo:Condition**
  - ppo:resourceAsSubject / resourceAsObject: to restrict the resource’s URI when it is either a subject or an object
  - ppo:classAsSubject / classAsObject: to restrict instances of classes that are either as a subject or an object
  - ppo:hasProperty: to restrict instances of properties
  - ppo:hasLiteral: to restrict particular values
Access Test Queries

- ppo:AccessSpace: defines SPARQL ASK queries that test a user’s information if it matches the graph pattern

Advantages:
- User’s don’t need to specify friends for each privacy preference
- Since users’ information change over time, the access space ensures that the correct type of users access the information
Access Control Privileges

- ppo:hasAccess: Defines the access privilege(s) which is granted within a privacy preference
  - Read / Write access to statements
  - Defined using Web Access Control (WAC) Vocabulary
Example:

A user wants to restrict a microblog post tagged with the concept of Linked Data to users that have a similar interest

```xml
<http://www.example.org/pp3> a ppo:PrivacyPreference;
ppo:appliesToResource <http://smob.me/user/xyz/post1>;
ppo:assignAccess acl:Read
ppo:hasCondition [ ppo:hasProperty tag:Tag;
ppo:resourceAsObject
<http://dbpedia.org/resource/Linked_Data> ];
ppo:hasAccessSpace [ ppo:hasAccessQuery
"ASK {
?x foaf:topic_interest
<http://dbpedia.org/resource/Linked_Data> }"
].
```
Applying the Privacy Preference Ontology

- A Privacy Preference Manager that provides users to specify privacy preferences for their FOAF files
- The privacy preference manager grants other users which information to access
Progress and Future Work

- **Progress so far:**
  - We developed the PPO
  - Currently, the Privacy Preference Manager is being developed

- **Future Work:**
  - To Extend the PPO to restrict actions
    - For instance: Allow messages sent from work colleagues and restrict any messages who are not work colleagues, if I am busy
  - To cater for conflicting privacy preferences
  - To investigate relationships with RDFS and OWL entailments