Defining Expressive Access Policies for Linked Data using the ODRL Ontology 2.0 *

Simon Steyskal, Axel Polleres
Vienna University of Economics and Business
Institute for Information Business
Welthandelsplatz 1, A-1020 Vienna
(simon.steyskal|axel.polleres)@wu.ac.at

ABSTRACT

Together with the latest efforts in publishing Linked (Open) Data, legal issues around publishing and consuming such data are gaining increased interest. Particular areas of interest include (i) how to define more expressive access policies which go beyond common licenses, (ii) how to introduce pricing models for online datasets (for non-open data) and (iii) how to realize (i)+(ii) while providing descriptions of respective meta data that is both human readable and machine processable. In this paper, we show based on different examples that the Open Digital Rights Language (ODRL) Ontology 2.0 is able to address all previous mentioned issues, i.e. is suitable to express a large variety of different access policies for Linked Data. By defining policies as ODRL in RDF we aim for (i) higher flexibility and simplicity in usage, (ii) machine/human readability and (iii) fine-grained policy expressions for Linked (Open) Data.

Categories and Subject Descriptors

H.3.5 [**Information Storage and Retrieval**]: Online Information Services

Keywords

linked data, license, digital rights, odrl, access policies

1. INTRODUCTION

Publishing data on the Web usually involves more than just making it accessible for the public. Following the design principles for publishing Linked Data proposed by Tim Berners-Lee [1] is a first step towards the publication of a well structured Linked Dataset but, unfortunately, those principles offer no guidelines for defining fine-grained and expressive meta data to which purposes and under which conditions such data may be used, such as e.g for defining access policies. The openness of the Web and the lack of standard mechanisms to define (and enforce) expressive access restrictions in a way that is both human readable and machine processable, might in fact prevent data owners from publishing their datasets [5]. Another important aspect is (monetary) cost caused by gathering and preparing the data. Despite the general trend towards openness and transparent availability of date, if there is no possibility to regain some of the expenses made during creating and curating a dataset, data owners might not see any benefit from publishing it.

In order to address these issues and to demonstrate data owners a possibility to easily define expressive access policies for Linked Data using a common rights expression language, in this short paper we look into the suitability of the Open Digital Rights Language 2.0 Ontology[1] as a candidate for achieving this task, by discussing example scenarios with their respective policies and their realization using the ODRL vocabulary. While previous work has already hinted on the possibility of using ODRL for expressing Linked Data licences [2]

The remainder of the paper is structured as follows: in Section 2 we give an introduction into ODRL and investigate constructs which are particularly important for the domain of Linked Data, Section 3 underpins our choice of using ODRL to express access policies for Linked Data by providing a set of use cases and their ODRL representations. In Section 4 we discuss related work in the field of access policies and restrictions for Linked Data before we conclude (Section 5).

2. OPEN DIGITAL RIGHTS LANGUAGE 2.0

The Open Digital Rights Language (ODRL) was invented to provide an open standard for defining policy expressions for digital content and media. The ODRL Core Model (cf. Figure1) contains all major components of an ODRL policy expression.

In the following we will discuss the different main components in more detail and will especially only focus on parts of the ODRL vocabulary which might be relevant for the domain of Linked (Open) Data.

Policy A `Policy` is the central entity that forms ODRL policy expressions. It can refer to `Permissions` and

*Simon Steyskal has been partially funded by the Vienna Science and Technology Fund (WWTF) through project ICT12-015.

SEM '14 September 04 - 05 2014, Leipzig, AA, Germany
ACM 978-1-4503-2927-9/14/09 $15.00.
http://dx.doi.org/10.1145/2660517.2660530

[1] `http://www.w3.org/community/odrl/`

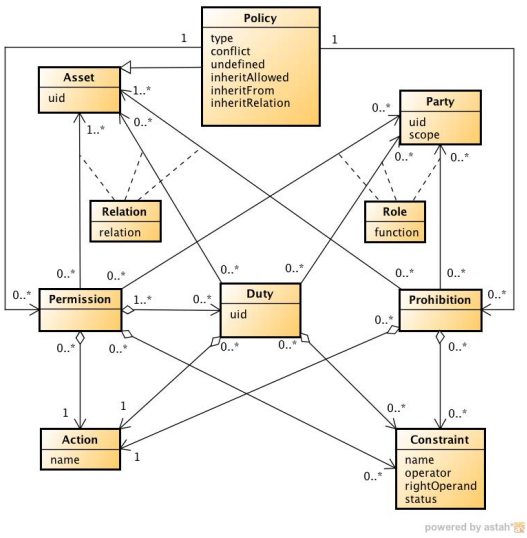

Figure 1: ODRL Core Model Version 2.0 taken from [6]

`Prohibitions` which hold for that `Policy`; policies be further distinguished into several subtypes, namely:

(i) An `Agreement` is a policy expression, which represent a formal contract or license containing the involved `Parties` and the respective terms of usage. In contrast to general license definitions, which are realized as `Sets`, an `Agreement` explicitly defines the `Parties` amongst the access conditions (i.e. `Permissions` and `Prohibitions`) were stipulated.

(ii) An `Offer` is used if an `Asset` owner wants to represent (e.g., as advertisement) possible access conditions which are usually linked to `Duties` in absence of a concrete `Party` consuming the `Asset`. (e.g. *Access to perform ASK Queries on a specific dataset is granted in exchange for a payment of 400 Euro.* or cf. Listing 3)

(iii) `Set`: if general terms of usage shall be defined which do not have any `Constraints` or `Duties` attached, `Sets` can be used. This type of `Policy` can also be used to represent common Open Data licenses such as Creative Commons (CC)[2] licenses.

(iv) A `Ticket` is a policy expression that stipulates the terms of usage between an `Asset` owner and any Party who currently holds such `Ticket` in possession.

Asset An `Asset` is the entity whose terms of usage are restricted by its surrounding policy expression. In the domain of Linked (Open) Data an `Asset` is usually a dataset or parts of a dataset.

Party A `Party` can be distinguished into `Assigner` (the party who proposes the policy statements) and a `Assignee` (the one who receives the policy statements).

Permission A `Permission` allows that a specific `Action` is executed on a particular `Asset` (e.g. *Read data from dataset <dataset1>.*). In addition,

(i) `Constraints` restrict that `Permission` and define it in more detail (e.g. *Read data from dataset <dataset1> until December* 31^{st}),

(ii) `Parties` can be defined between whom the policy is stipulated (e.g. *Company XYZ allows user Alice to read data from dataset <dataset1> until December* 31^{st}), as well as

(iii) `Duties` which must be fulfilled beforehand in order for the `Permission` to become valid (e.g. *Company XYZ allows user Alice to read data from dataset <dataset1> until December* 31^{st}*, if Alice pays 400EUR.*),

can be attached.

Prohibition In contrast to `Permissions`, `Prohibitions` are used to forbid specific `Actions` on an `Asset` (e.g. *In general, no one is allowed to read data from dataset <dataset1>.*) and cannot refer to `Duties`.

Duty As already exmplified above, a `Duty` defines a certain `Action` which has to be executed by a potential `Assignee` for the `Permission` to become valid.

Action `Actions` are operations which a potential `Assignee` is allowed (if related to a `Permission`), is prohibited (if related to a `Prohibition`) or has (if related to a `Duty`) to perform. We have investigated different actions which can be used to describe common actions taken within a Linked Data scenario, namely:[3]

aggregate *"The act of using the asset (or parts of it) as part of a composite collection."* In the domain of Linked Data, `aggregate`s can be used to express the action of querying different datasets

[2] `http://creativecommons.org/ns`

[3] For a list of available actions, we refer readers to `http://www.w3.org/community/odrl/two/vocab/`

and aggregate the retrieved data (e.g. an asset owner can permit the querying of its dataset but only if the results are not aggregated with those from other data sources).

read *"The act of obtaining data from the asset."* Among the various different `Actions` defined withing the ODRL vocabulary `read` does best fit the intended meaning of a SPARQL query or retrieving (parts of) a dataset, in general.

write *"The act of writing to the asset."* Thinking in the scope of querying Linked Data resources, `write` can be used to represent SPARQL INSERT, and may be used for SPARQL UPDATE queries, although the latter case does not fit the intended semantics of `write` perfectly since `write` only specifies the action of *writing to* and not *deleting from* the asset.

delete *"The act of permanently removing the asset."* `delete` can be used to express SPARQL DELETE and (partially) SPARQL UPDATE, but again without having the same intended semantics, not capturing inserts. While `delete` specifies the complete deletion of an `Asset`, SPARQL DELETE and (partially) SPARQL UPDATE usually only delete parts from an `Asset`.

nextPolicy Generally speaking, this `Action` is used to refer to policy restrictions which have to hold for derived `Assets` of the original `Asset`.

Constraint With `Constraints` it is possible to restrict and limit the scope of `Permissions, Prohibitions` and `Duties`, using a simple mathematical structure with two operands and one operator (e.g. *"The number of query requests (`operand`) must be less or equal (`operator`) than 100 (`operand`)"*. Such constraints could – again in the context of SPARQL – be used to advertise restrictions on the usage of services such as SPARQL endpoints, e.g. using ODRL as part of a SPARQL 1.1 service description [14].

3. EXPRESSING ACCESS POLICIES FOR LINKED DATA IN ODRL 2.0

Although there are numerous possible scenarios where access restrictions play an important role, we have identified six common use-cases which should be expressible and which we will realize with ODRL.

Restricting access to specific parties and datasets Maybe one of the most common and basic requirements for any rights expression language is the possibility to restrict access to specific users and/or restrict access to specific assets. A possible example is shown in the example underneath, where user `:alice`is granted the permission to read only datasets `dataset1` and `dataset2`.

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix : <http://www.example.com/> .

<listing1> a odrl:Permission;
     odrl:action odrl:read;
     odrl:target
         </dataset1>, </dataset2>;
     odrl:assigner :owner ;
     odrl:assignee :alice .
```

Limiting number of request A more advanced policy is exemplified in the following listing. It expresses an access policy which restricts the number of allowed read requests for the purpose of `sp:Ask` (i.e. performing ASK queries) to 100 and furthermore stores its current status. Such a scenario is especially feasible in combination with payment duties, where a party gets the permission to perform a number of requests if it pays a certain amount for it.

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix spin: <http://spinrdf.org/sp/> .

<listing2> a odrl:Permission;
     odrl:action odrl:read;
     odrl:target </dataset>;
     odrl:constraint [
          a odrl:Constraint;
         odrl:purpose sp:Ask;
         odrl:operator odrl:lteq;
         odrl:count "100"^^xsd:integer;
         odrl:status "42"^^xsd:integer ] .
```

Allowing access only in specific time windows It is also possible to allow (or forbid) access within specific time windows only as illustrated in the listing underneath. Again, this is especially useful in scenarios where data owners want to restrict access to data or services based on server workload or payment issues (e.g. more popular time windows are more expensive).

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix spin: <http://spinrdf.org/sp/> .

<listing3> a odrl:Permission;
     odrl:action odrl:read;
     odrl:target </dataset>;
     odrl:constraint [
         a odrl:Constraint;
         odrl:purpose sp:Ask;
         odrl:operator odrl:lteq ;
         odrl:dateTime
             "2016-12-31"^^xsd:date ] .
```

Representing licenses As already discussed, ODRL can also be used to represent common licenses in RDF as it includes all necessary constructs to do so.

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix : <http://www.example.com> .

% CC-BY-NC-SA
<listing4> a odrl:Set;
     odrl:permission odrl:reproduce ;
     odrl:permission odrl:distribute ;
     odrl:permission odrl:derive ;
     odrl:duty odrl:attribution ;
     odrl:duty odrl:attachPolicy ;
     odrl:duty odrl:shareAlike ;
     odrl:prohibiton odrl:commercialize .
```

Restricting data-reuse policy An important – an so far under-researched in terms of technical solutions [9] – aspect of publishing data on the web and especially in combination with restricted data is the definition of data re-use policies. Although a data owner might have granted a party to use its data, the owner usually does not want to pass this permission to downstream

customers. ODRL offers the possibility to explicitly define such re-use policies.

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix : <http://www.example.com/> .

<listing5> a odrl:Permission;
     odrl:action odrl:distribute;
     odrl:target </dataset>;
     odrl:duty [
          a odrl:Duty ;
          odrl:action odrl:nextPolicy ;
          odrl:target :newPolicy ] .

:newPolicy a odrl:Set ;
     odrl:permission [
          a odrl:Permission ;
          odrl:action odrl:display ;
          odrl:target </dataset> ] .
```

Introducing payment duties As already mentioned above, the possibility to define payment duties which have to be fulfilled in order get a certain permission is a crucial part for our Linked Data scenario. In ODRL, such payment duties can be easily defined and assigned to any permission or duty as exemplified in the listing underneath.

```
@prefix odrl: <http://w3.org/ns/odrl/2/> .
@prefix gr: <http://purl.org/goodrel/v1#>.

<listing6> a odrl:Permission;
     odrl:action odrl:read;
     odrl:target </dataset>;
     odrl:duty [
          a odrl:Duty ;
          odrl:action odrl:pay ;
          odrl:target [
             a gr:UnitPriceSpecification;
             gr:hasCurrencyValue
                   "400"^^xsd:float;
             gr:hasCurrency
                   "EUR"^^xsd:string
          ] .
     ] .
```

4. RELATED WORK

To the best of our knowledge, we are the first to suggest the ODRL vocabulary for defining more expressive access policies in RDF for Linked Data. Cabrio et al. [2] use ODRL to represent licenses in their recently proposed natural language approach for automatically generating RDF licenses. In addition to represent licenses as ODRL policies, we use ODRL to express more detailed access restrictions. Other approaches which use ontologies to define access policies are the *Privacy Preference Ontology (PPO)* [12] and the *SHI3LD Model* [4, 5]. Both of them define their own ontology to be used for defining access policies and use `SPARQL ASK` queries to verify them. Our approach differs from theirs, since we reused an already established standard rights expression language rather than defining our own one for modeling access policies.

Other approaches such as Mühleisen et al. [8], which propose a SWRL-based strategy to define access policies, or Villata et al. [13] and Rodriguez et al. [11] which focus on investigating and representing license definitions for Linked Data, address only parts of possible scenarios where access restriction is necessary, but could be relevant for our next step, which is formalizing ODRL policies for linked data in a machine-processable way. To this end we shall check other earlier machine-readable and -processable formalizations of policies such as PROTUNE [3], REI [7], or XACML [10].

5. CONCLUSIONS

In this paper we proposed ODRL as suitable for expressing more fine-grained access policies for Linked Data by investigating the core concepts of ODRL itself and providing a number of listings to underpin our proposition. While ODRL was initially intended to be used to define an open standard for policy expressions for digital media, we proved that it is also suitable for expressing access policies within the domain of Linked Data.

Future work will include the development of a framework, capable of using and checking access policies defined in ODRL and a more thoroughly evaluation of different access policy frameworks.

6. REFERENCES

[1] Tim Berners-Lee. Linked data design principles. `http://www.w3.org/DesignIssues/LinkedData.html`, June 2014.

[2] Elena Cabrio, Alessio Palmero Aprosio, and Serena Villata. These are your rights. In *The Semantic Web: Trends and Challenges*, pages 255–269. Springer, 2014.

[3] Juri Luca De Coi, Daniel Olmedilla, Piero A. Bonatti, and Luigi Sauro. Protune: A framework for semantic web policies. In Christian Bizer and Anupam Joshi, editors, *International Semantic Web Conference (Posters & Demos)*, volume 401 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.

[4] Luca Costabello, Serena Villata, Nicolas Delaforge, and Fabien Gandon. Shi3ld: an access control framework for the mobile web of data. In *Proceedings of the 23rd ACM conference on Hypertext and social media*, pages 311–312. ACM, 2012.

[5] Luca Costabello, Serena Villata, and Fabien Gandon. Context-aware access control for rdf graph stores. In *ECAI*, pages 282–287, 2012.

[6] ODRL Community Group. Odrl 2.0 core model. `http://www.w3.org/community/odrl/two/model/`, June 2014.

[7] Ryusuke Masuoka, Mohinder Chopra, Yannis Labrou, Zhexuan Song, Wei-lun Chen, Lalana Kagal, and Timothy Finin. Policy-based access control for task computing using rei. In *Policy Management for the Web Workshop*, pages 37–43, 2005.

[8] Hannes Muhleisen, Martin Kost, and Johann-Christoph Freytag. Swrl-based access policies for linked data. In *SPOT 2010 2nd Workshop on Trust and Privacy on the Social and Semantic Web, Heraklion, Greece*, 2010.

[9] Alexander Novotny and Sarah Spiekermann. Personal information markets and privacy: A new model to solve the controversy. In *11. Internationale Tagung Wirtschaftsinformatik*, page 102, 2013.

[10] Torsten Priebe, Wolfgang Dobmeier, and Nora Kamprath. Supporting attribute-based access control with ontologies. In *Availability, Reliability and Security, 2006. ARES 2006. The First International Conference on*, pages 8–pp. IEEE, 2006.

[11] Vıctor Rodrıguez-Doncel, Asunción Gómez-Pérez, and Nandana Mihindukulasooriya. Rights declaration in linked data. In *Proc. of the 3rd Int. Workshop on Consuming Linked Data*. Citeseer, 2013.

[12] Owen Sacco, Alexandre Passant, and Stefan Decker. An access control framework for the web of data. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on*, pages 456–463. IEEE, 2011.

[13] Serena Villata and Fabien Gandon. Licenses compatibility and composition in the web of data. In Juan Sequeda, Andreas Harth, and Olaf Hartig, editors, *COLD*, volume 905 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2012.

[14] Gregory Williams. SPARQL 1.1 Service Description. W3C recommendation, W3C, March 2013. `http://www.w3.org/TR/sparql11-service-description/`.