SciDataCon 2016

Conference paper

Agreeing about agreements: modelling social contracts, people and data

Summary

One of the well-established methods used to ease data sharing between organisations and even teams within organisations is to use standards for data structure, metadata and interfaces. Standards are a form of agreement, as are MoUs, charters, deeds, licences, rules of the road and even the definitions for words. Many of these other sorts of agreements are also important for data sharing communities too.

In this paper we look to improve the efficiency of dealing with different forms of agreement within a data sharing scenario by presenting a prototype agreements ontology which models agreements themselves as 'things' and the relationships between them and between them and data and them and agents. Having an agreements ontology allows us to start automating tasks that require knowledge of them. This may take the form of data repositories that can make intelligent choices about how to deliver or withhold data without human intervention. We position this ontology as a 'middle' ontology, that is one which specializes well-known, abstract, upper ontologies and is able to be used fairly widely but is expected to be used in particular contexts in conjunction with detailed, domain-specific, lower ontologies. We have relied on existing agent, data manipulation, and metadata ontologies where possible and as such we specialise the ORG and FOAF ontologies, the PROV ontology and DCAT and ODRS ontologies for those areas respectively.

This paper and ontology supports work that we report elsewhere at SciDataCon2016, namely *The Role of Social Architecture in Information Infrastructure* (**Box & Lemon**) and *Describing and Automating Requirements within Licenses and their Resolutions* (**Car & Stenson**).

Background

Australia has built a series of large inter-agency and sometimes intergovernmental information infrastructures over the last decade. These include initiatives to bring together observational and modelled data bout the Great Barrier Reef (eReefs), store all of Australia's water supply organisation's accounts (Water Regulations), hold data on Australia's living species (Atlas of Living Australia) and share terrestrial ecosystems' data (TERN)[1]. These actives are undertaken within a complex legislative, policy, organizational and social context. Furthermore, authority structures are established for each initiative to govern communities. A range of agreements, required to facilitate data sharing are created through these structures. Agreements provide the rules of the game for those participating in collaborative activities and range from policy and standards through data licensing to repository policy and inter-organisational Memoranda of Understanding (Box and Lemon 2015). In all circumstances, the interplay between the various forms of agreement created within initiatives and those that exist independently of them, is not explicitly described either in relation to the data sharing that the agreements facilitate, or the agents (organisations and people) that generate and

subscribe to them. This opaque situation creates confusion, conflict and cost for those trying to share and access data. The explicit modelling and subsequent declaration of agreements and their relation to data and agents can assist in streamlining data sharing.

The Agreements Ontology

We have created an Agreements Ontology (AGR-O) presented at http://promsns.org/def/agr which is also the AGR-O namespace and we suggest `agr` for its prefix. We have chosen OWL[2] as our knowledge representation language due to its ease of use with a range of existing standards such as the Resource Description Format (RDF[2]) for its data format, the SPARQL[2] query language for querying data delivered using it and the existence of many other OWL ontologies that cover much of the conceptual ground relating to data, agents and some forms of agreements (e.g. licenses) from which we can inherit.

An agreement as an *entity*

Modelling in OWL, an agreement (`agr:Agreement`) is, by definition, a subclass of `owl:Thing`[3] and beyond existence/non-existence we classify it as an *entity*, as opposed to an *agent* or an *agent*, distinctions made by the upper ontology PROV-O (Lebo *et al.* 2013). PROV-O has a concept of a thing (`prov:Entity`) specialised as a `prov:Plan` which instructs agents (`prov:Agent`) in some course of action. Since agreements lead to action, we create a subclass of `prov:Plan` for them. This hierarchy is given in **Figure 1A**.

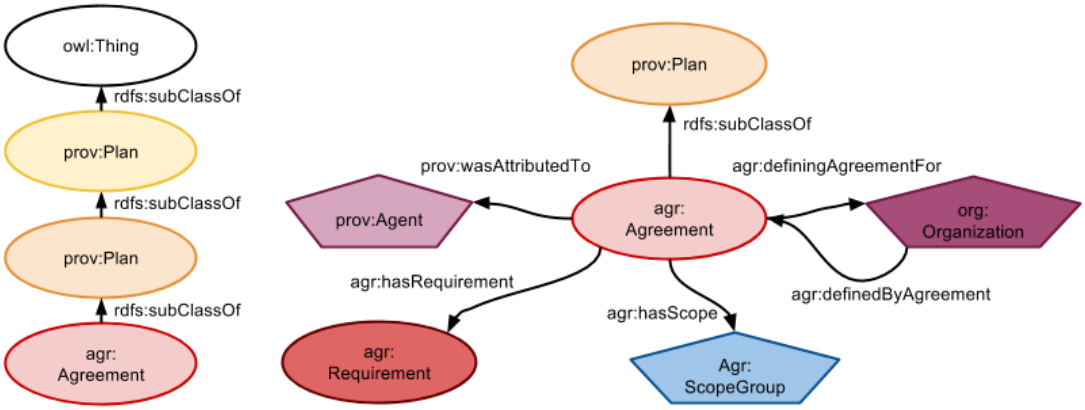

Figure 1 A (left): The `agr:Agreement` class in relation to PRV-O & OWL classes; **B** (right): Basic properties and classes immediately related to `agr:Agreement`.

Making agreements

Agents make agreements and, modelled agreements as a subclass of `prov:Entity`, via `prov:Plan`, means PROV-O entity/agent relationships apply e.g., agreements can be created (`prov:generated`) or destroyed (`prov:invalidated`) by `prov:Agents` via activities (`prov:Activity`) they undertake (Lebo *et al.* 2013).

How agreements apply to agents

We think of agreements applying to parties which we can model as `prov:Agent`s. We link agreements (`agr:Agreement`) via an object property of `agr:hasScope` (inverse: `agr:scopeOf`) to a `agr:ScopeGroup` class which is a subclass of `foaf:Group` and thus something that a `foaf:Agent`/`prov:Agent` can be a member of. Since agreements may affect some agents differently to others, we allow agents to have qualified membership of the

`agr:ScopeGroup` following the pattern of PROV-O's qualification properties. This is a role-based extension to ORG-O's memberships. See the ontology specification for examples.

Other Agent relationships

There are many other agreement/agent relationships possible through PROV-O and ORG-O combinations and we define some with shortcuts for convenience, e.g. agent creation. A charter may establish a company or an organisation's members agree to create a committee. For this we define the agreement class object property of `agr:definingAgreementFor` (inverse: `agr:definedByAgreement`) with a range value of a `org:Organization`. This is a shortcut to PROV-O modelling whereby a `prov:Agent` is viewed as a `prov:Entity` and created (`prov:wasGeneratedBy`) by a `prov:Activity` which could take in (`prov:used`) an `agr:Agreement` as an input (since an `agr:Agreement` is a `prov:Plan`). These basic properties of an `agr:Agreement` are given in **Figure 1B**.

How agreements affect agents

The Creative Commons Rights Expression Language (CCREL)[4] has the notion of a requirement which is an "an action that may or may not be requested of you" as a result of a license. We generalise this to applying to agreements (linked via `agr:hasRequirement`), thus the effect of agreements on agents is described through a series of individual requirements with our goal being for maximal requirement reuse among agreements to help with compliance simplification. As per our qualified membership of agents to an agreement's scope group, we allow qualification of how a requirement compels an agent to action, taking direction from **Bradner (1997)**. This is in contrast to the CCREL modelling which classes requirements and prohibitions separately: in `agr`, all the required actions of an agreement are of type `agr:Requirement`. As per **Car & Stenson (2015)** we propose a mechanism for the automated recognition of compliance with requirements by the generation of `agr:RequirementResolution` objects by agents seeking to demonstrate compliance.

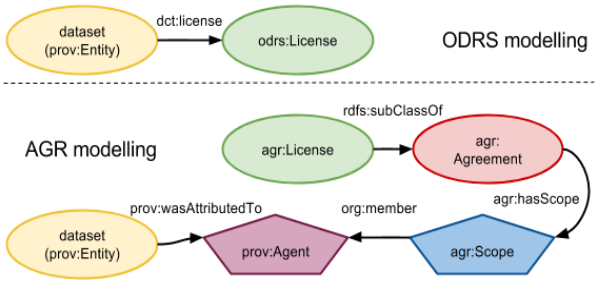

Figure 2: Relating datasets to agreements via agents in AGR-O compared with the ODRS[5].

No agreements hierarchy

Initially, we modelled a hierarchy of different agreement types, such as contracts, license, Memoranda of Understanding, standards etc. however we now believe that no sensible hierarchy can be created for generic use. We prefer to determine the relative order of agreements in confined situations based on their applicability to agents and via a scope group with qualifications and their requirements of those agents via qualified requirements.

Applying agreements to data

We are used to linking data directly to some forms of agreement, for example licenses. In this ontology, agreement/dataset linking is accomplished via agreement/agent and agent/dataset linking. Remodelling the Open Data Rights Statement's[5] dataset/licence relationship according to this paradigm is shown in **Figure 2**. This retains the social component of agreements, that they directly affect agents and their actions, not things.

Examples

We present dummy ontology instances in the examples section of http://promsns.org/def/agr to demonstrate the following scenarios: 1. Discovering the agreements pertaining to a particular dataset; 2. Discovering datasets based on the agreements pertaining to them; 3. Discovering the requirements pertaining to an agent via the agreements they are within the scope of; 4. Resolving the requirements of conflicting agreements (licence v. MoU); and 5. Demonstration of the satisfaction of agreements for automated agreement adherence.

Conclusions

This ontology is a start in allowing us to deal with agreements in a systematic and sometimes automatic way. Its primary goal is to facilitate machine readable access to and understanding of data via the social agreements that pertain to it. This work is open for comment and we expect and hope for extensions to it and specializations of it.

Competing Interests

The authors declare that they have no competing interests.

Notes

1. eReefs: https://research.csiro.au/ereefs/, Water Regulations: http://www.bom.gov.au/water/regulations/aboutRegs.shtml, Atlas of Living Australia: http://www.ala.org.au/, TERN's data project: http://tern.org.au/Data-pg22386.html
2. OWL: https://www.w3.org/standards/techs/owl, RDF: https://www.w3.org/RDF/, SPARQL: https://www.w3.org/TR/sparql11-overview/
3. OWL is a mathematical set theory-based language and `owl:Thing` is the set of all individuals, meaning everything in OWL is a subclass of `owl:Thing`.
4. http://creativecommons.org/ns
5. http://schema.theodi.org/odrs/

References

Box P. and **Lemon D.** 2015 The Role of Social Architecture in Information Infrastructure: Report for the National Environmental Information Infrastructure (NEII). CSIRO, Australia.

Bradner, S.1997 Key words for use in RFCs to Indicate Requirement Levels. March 1997. Internet RFC 2119. Online http://www.ietf.org/rfc/rfc2119.txt [Last accessed 30 May 2016].

Car, N.J. and **Stenson, M.P.** 2015. Complex licence requirements for the Bioregional Assessments Programme managed by provenance. Proceedings of MODSIM2015, 21st International Congress on Modelling and Simulation. MSSANZ, December 2015, pp. 613–619. ISBN: 978-0-9872143-5-5. http://mssanz.org.au/modsim2015/C4/car.pdf

Lebo, T., **Sahoo, S.**, & **McGuinness, D.** (eds.) 2013 PROV-O: The PROV Ontology. W3C recommendation. Online http://www.w3.org/TR/prov-o/ [Last accessed 18 May 2016].