



Leibniz-Institut für
Astrophysik Potsdam



Provenance Data Model

Extra Meeting

InterOp Sesto, June 2015

Kristin Riebe, GAVO

Provenance DM Extra Meeting

- Instrument characteristics + ambient conditions
 - Should be modeled elsewhere, as entities
 - Use provenance data model only for *relations* between data/activities/agents

Provenance DM Extra Meeting

- Prototypes: yes or no?
 - Is W3C enough? Too abstract?
 - Or include prototypes, at least “Method” for “Activities”?
 - Using known methods and data-types could make it more suited for astronomical use cases; or put this information into a vocabulary/thesaurus
 - Could use something like Prov-Templates
 - What's the experience of CTA-people?

Provenance DM Extra Meeting

- RoleMap classes
 - Do we need them explicitly? For each annotated relation?
 - Or are annotated relations (graphs) enough?
 - Maybe extra classes needed for VO-DML?

 - Extra classes are useful, if entities and activities of different types are stored as the same object-type, i.e. in the same database table. Entities can take different roles for different activities.
 - If there are subtypes used for each type of activity, this is probably not needed.

Provenance DM Extra Meeting

- Other relations/types from W3C
 - Which ones could be useful, which ones not?

Types and Relation in W3C (Table 6)

Role in IVOA DM:

- needed
- maybe
- not needed?

	Entity	<code>entity(id, [attr1=val1, ...])</code>	Component 1: Entities/Activities
	Activity	<code>activity(id, st, et, [attr1=val1, ...])</code>	
	Generation	<code>wasGeneratedBy(id;e,a,t,attrs)</code>	
	Usage	<code>used(id;a,e,t,attrs)</code>	
	Communication	<code>wasInformedBy(id;a2,a1,attrs)</code>	
	Start	<code>wasStartedBy(id;a2,e,a1,t,attrs)</code>	
	End	<code>wasEndedBy(id;a2,e,a1,t,attrs)</code>	Component 2: Derivations
	Invalidation	<code>wasInvalidatedBy(id;e,a,t,attrs)</code>	
	Derivation	<code>wasDerivedFrom(id; e2, e1, a, g2, u1, attrs)</code>	
	Revision	<code>... prov:type='prov:Revision' ...</code>	
	Quotation	<code>... prov:type='prov:Quotation' ...</code>	Component 3: Agents, Responsibility, Influence
	Primary Source	<code>... prov:type='prov:PrimarySource' ...</code>	
	Agent	<code>agent(id, [attr1=val1, ...])</code>	
	Attribution	<code>wasAttributedTo(id;e,ag,attr)</code>	
	Association	<code>wasAssociatedWith(id;a,ag,pl,attrs)</code>	Component 4: Bundles
	Delegation	<code>actedOnBehalfOf(id;ag2,ag1,a,attrs)</code>	
	Plan	<code>... prov:type='prov:Plan' ...</code>	
	Person	<code>... prov:type='prov:Person' ...</code>	
	Organization	<code>... prov:type='prov:Organization' ...</code>	
	SoftwareAgent	<code>... prov:type='prov:SoftwareAgent' ...</code>	
	Influence	<code>wasInfluencedBy(id;e2,e1,attrs)</code>	Component 5: Alternate
	Alternate	<code>alternateOf(alt1, alt2)</code>	
	Specialization	<code>specializationOf(infra, supra)</code>	Component 6: Collections
	Collection	<code>... prov:type='prov:Collection' ...</code>	
	EmptyCollection	<code>... prov:type='prov:EmptyCollection' ...</code>	
	Membership	<code>hadMember(c,e)</code>	

Provenance DM Extra Meeting

- Application area
 - Just for observations (ObsProvenanceDM)?
 - From discussion with Mathieu:
 - Also model scientific proposal submission/acceptance/scheduling?
 - Model UWS-submissions?

Provenance DM Extra Meeting

- Accessibility of intermediate data products
 - Ignore non-public data? Flag them?
 - Provide only their header/metadata, more details on request?
 - Different provenance “views” for public/non-public?

Provenance DM Extra Meeting

- Level of detail
 - How to allow more/less details?
 - “Plan” for summarizing e.g. pipeline, workflow in one entity
 - “Bundle” mechanism allows to have provenance of provenance

Provenance DM Extra Meeting

- Collections
 - Can/Should/Shouldn't we allow entities to belong to >1 collection?
 - Need then associations, or model this using a mapping class+composition+references

Provenance DM Extra Meeting

- Vocabulary lists
- Links to other data models
- Name spaces/prefix: what will they mean, where to define