



Leibniz-Institut für
Astrophysik Potsdam

Provenance: IVOA vs. W3C

2019-09-17, Groningen

Ole Streicher



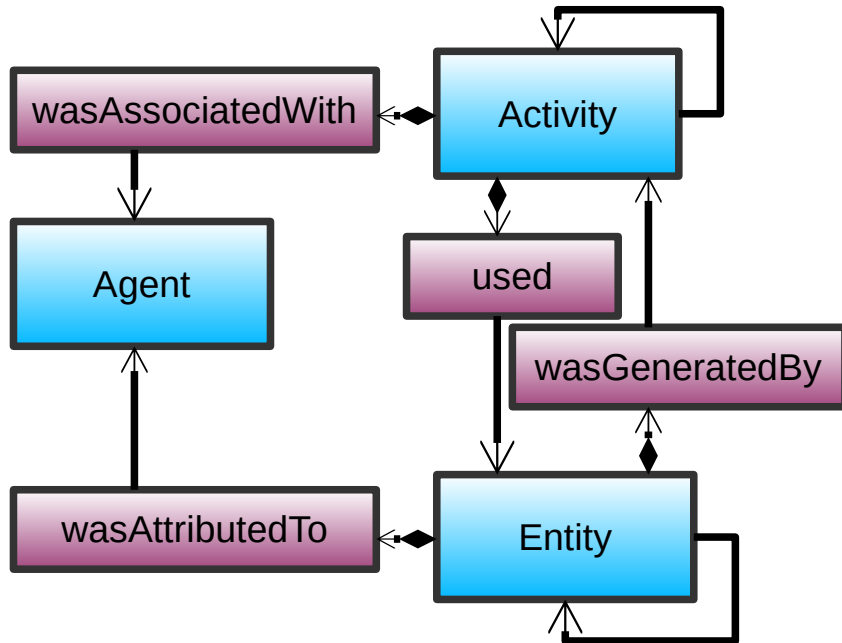
ProvSAP

- REST implementation of the IVOA Provenance model
- W3C compliant serializations
- Get provenance of a given Entity (data file) or Activity (observation, program run)
- Normal (backward) or forward provenance
- URL parameters control direction, depth, side paths (different, unfinished story)

W3C Provenance model

- Standard since 2013
- Relations between Activity, Entity, Agent
- Standard serializations: xml, json, rdf, prov-n
- Extension points:
 - prov:role
 - prov:type
 - Domain specific attributes
- Base of the IVOA Provenance standard
- But: different history → not fully compliant

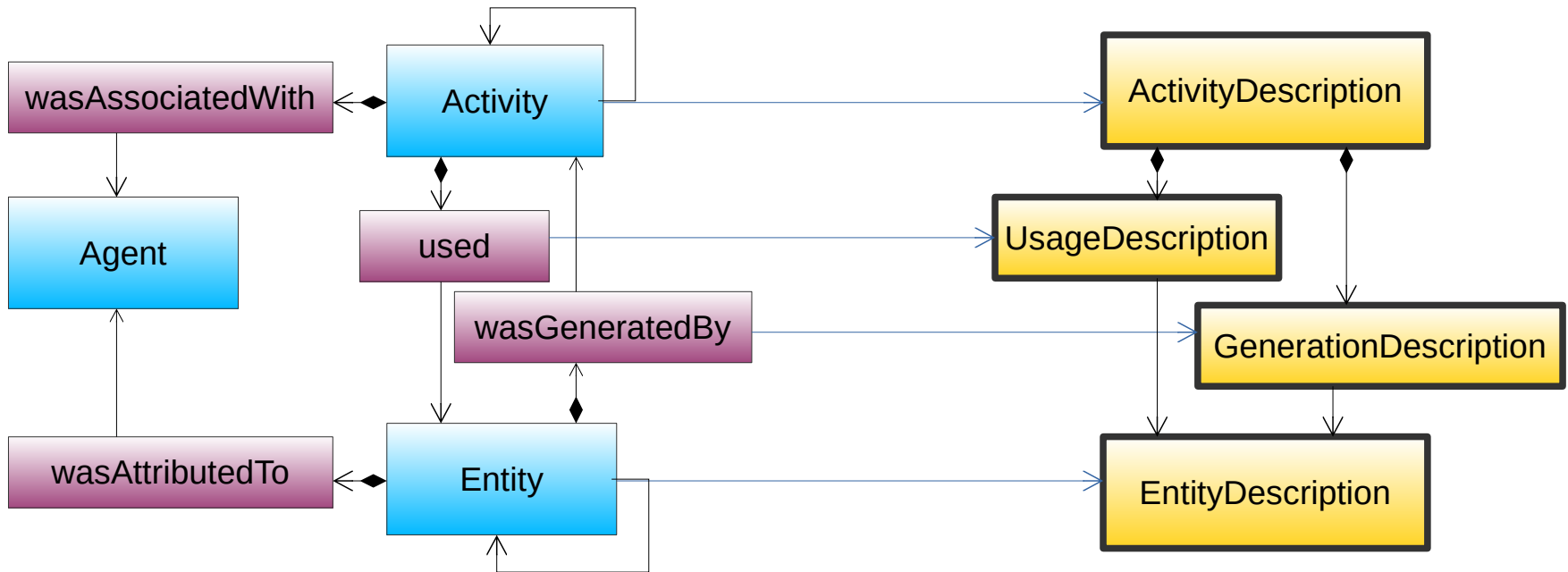
Provenance data model core classes



- 3 core classes:
 - Activity
 - Entity
 - Agent
- core relations:
 - Used
 - WasGeneratedBy
 - WasAttributedTo
 - WasAssociatedWith
- additional relations
 - wasDerivedFrom
 - wasInformedBy

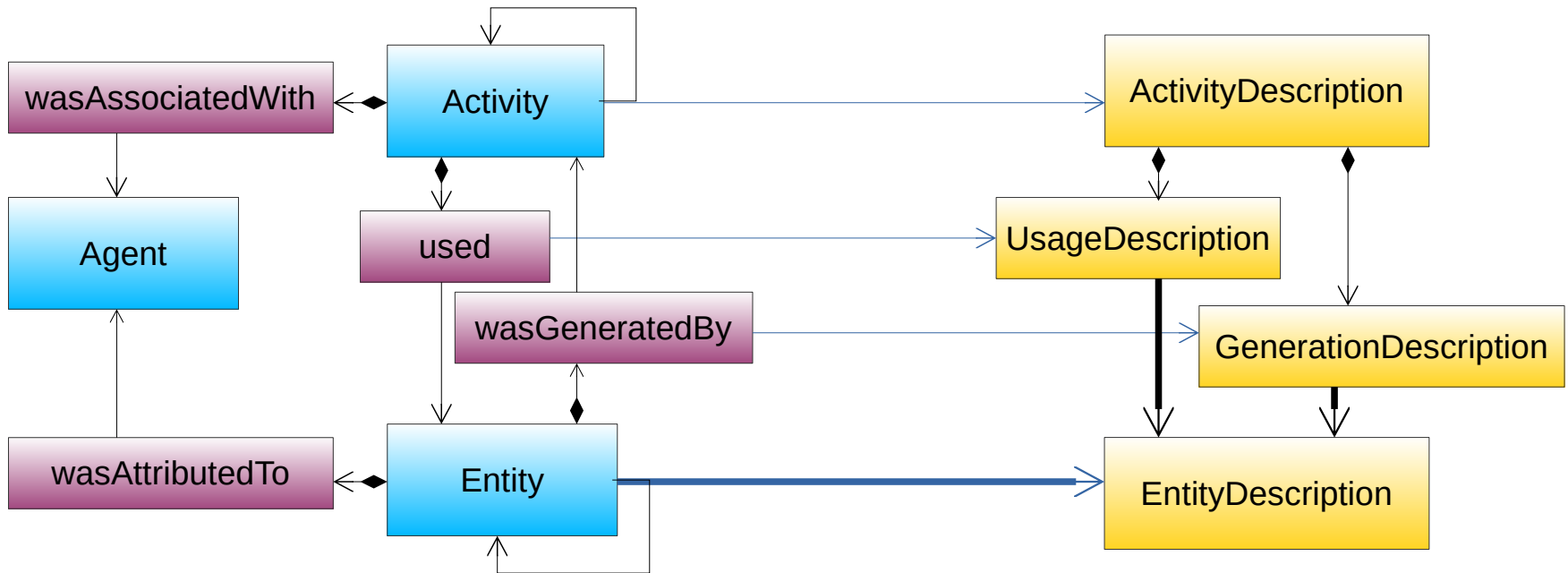
Description classes

- Not part of the W3C standard
- **Could be represented as Entities**
 - no Entity id
 - no provenance of these objects
- Alternative I: handle them outside of W3C Provenance
- Alternative II: Integrate contents into main classes



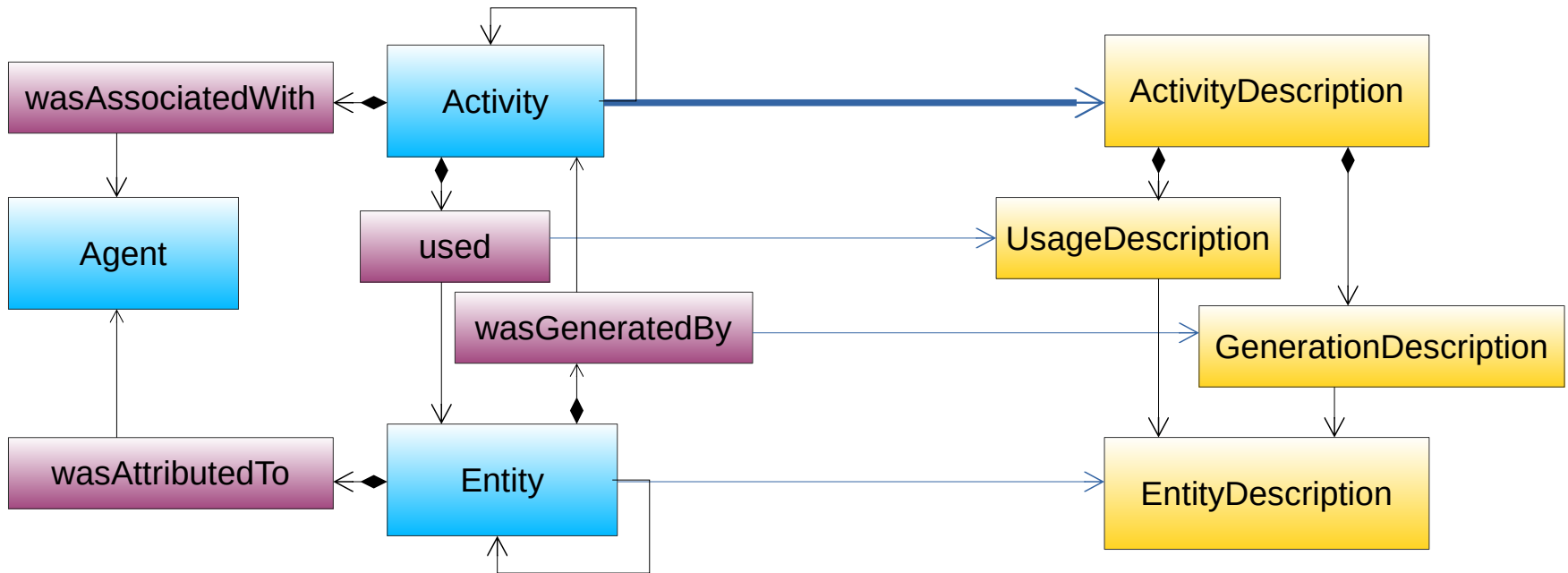
Relations to EntityDescription

- **Entity** → **EntityDescription**
- **UsageDescription, GenerationDescription** → **EntityDescription**
- EntityDescription may be created independently of other descriptions
- So, no prov:WasDerivedFrom
- modelled as simple link (attribute with the EntityDescription.id)



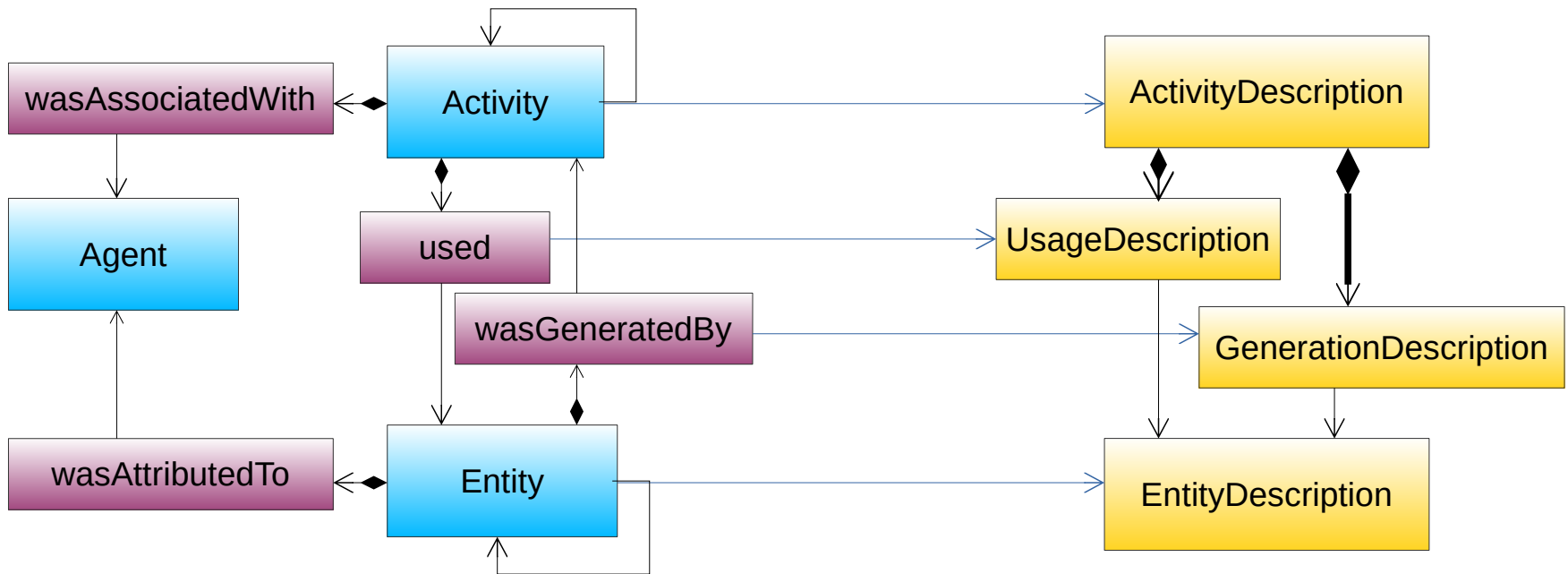
Relations to ActivityDescription

- **Activity** → **ActivityDescription**
- both are created independently
- therefore: not a `prov:WasInfluencedBy` relation
- Modelled as simple link (attribute with `ActivityDescription.id`)



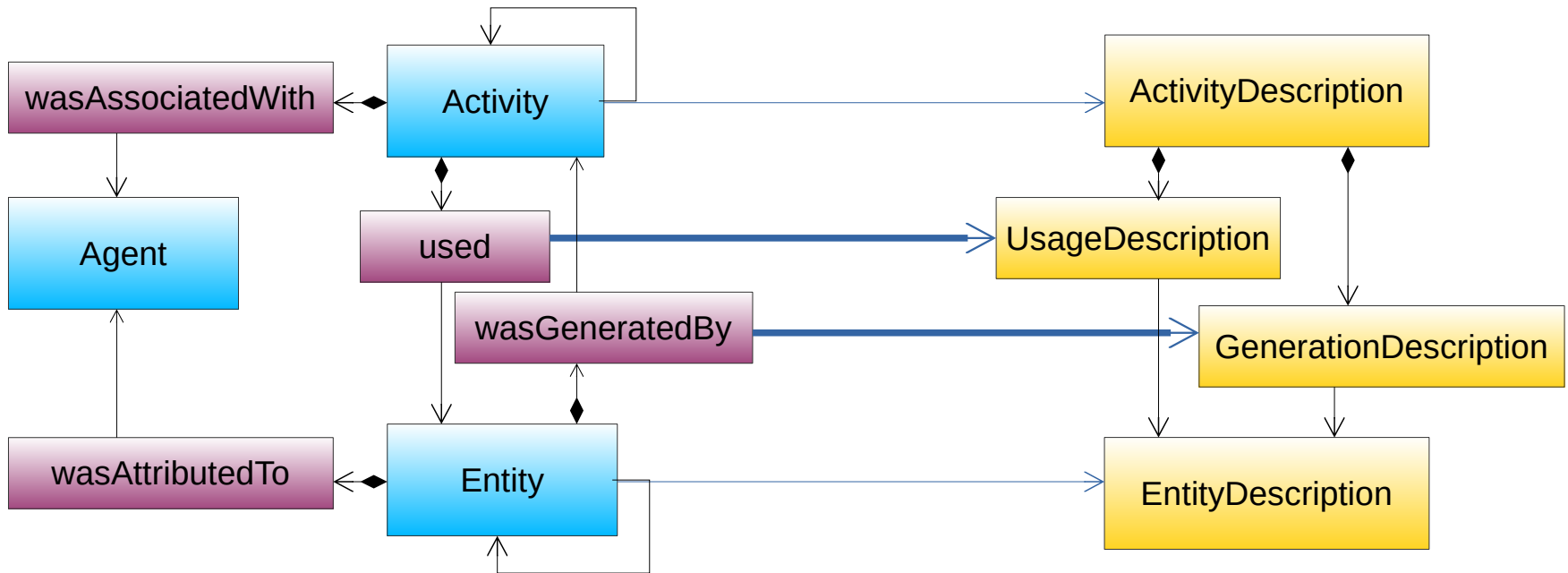
Relations from ActivityDescription

- **ActivityDescription** → **UsageDescription**
- **ActivityDescription** → **GenerationDescription**
- `prov:hadMember?` Then **ActivityDescription** would be a `prov:Collection`
- `prov:specializationOf?`



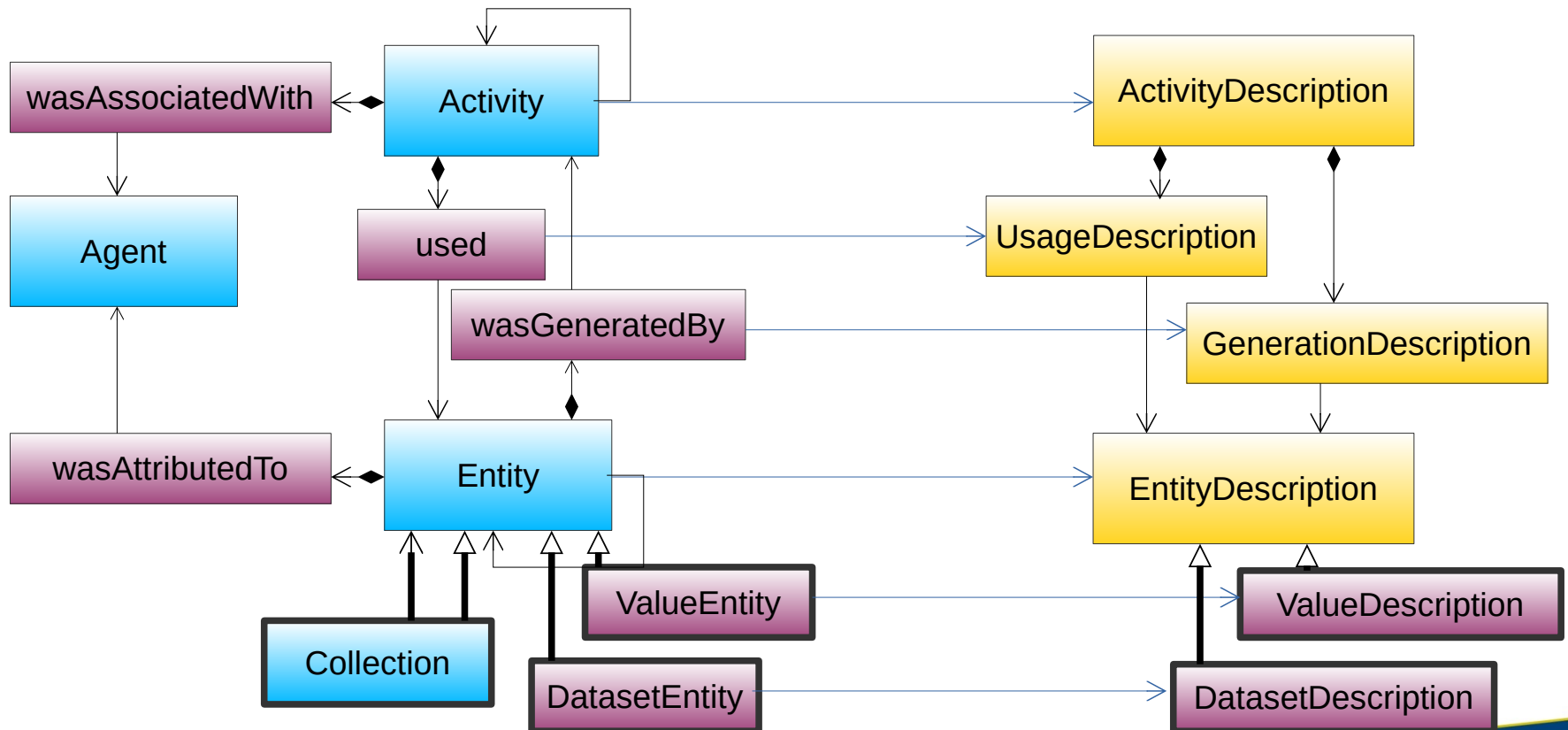
Relations to Usage/GenerationDescriptions

- Used and WasGeneratedBy are relations
- No relation between a relation and an Entity
- Possible solution:
 - Use prov:role to link them with id of the description
 - role name then becomes prov:label in the descriptions

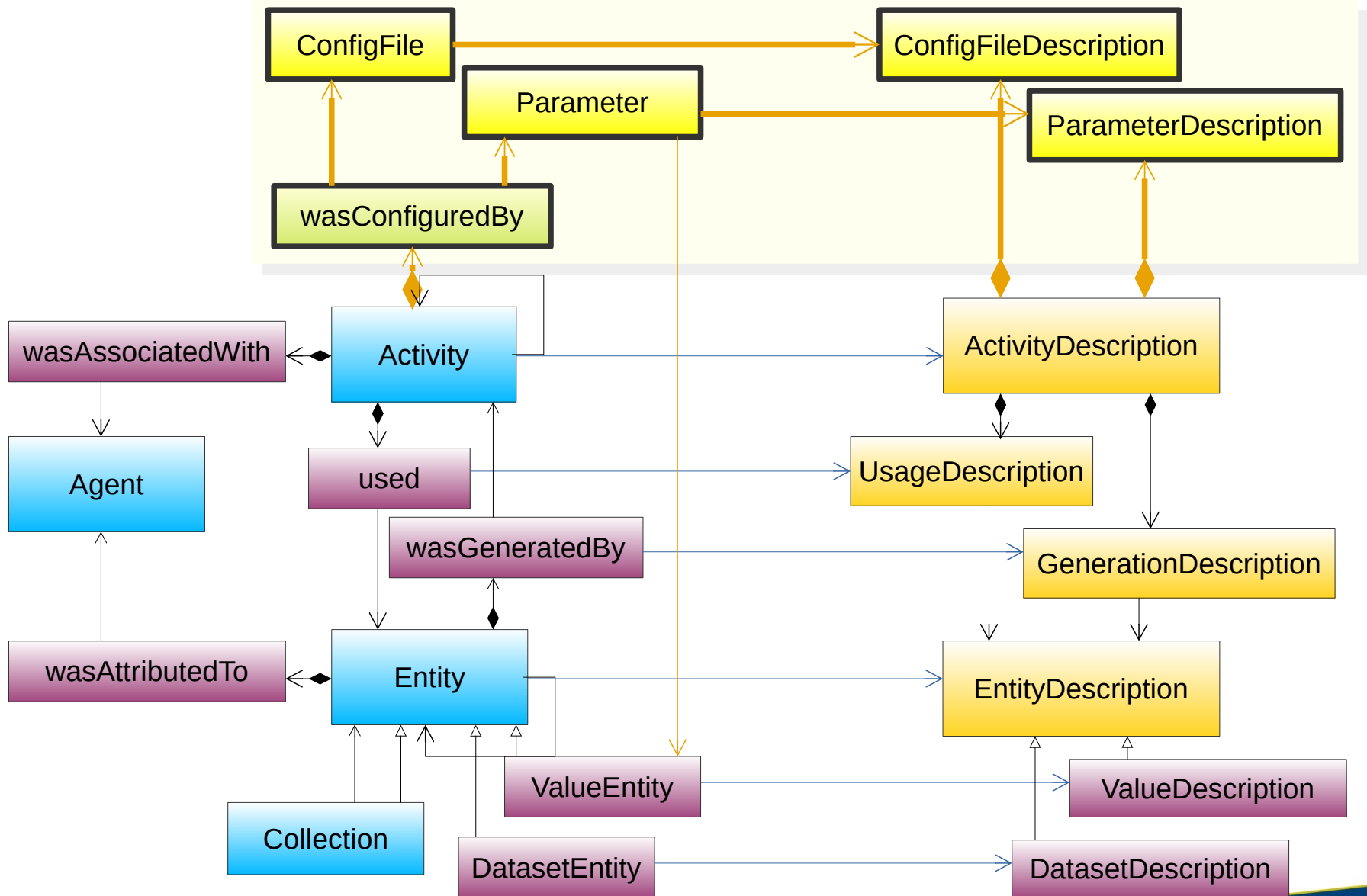


Specialized entities

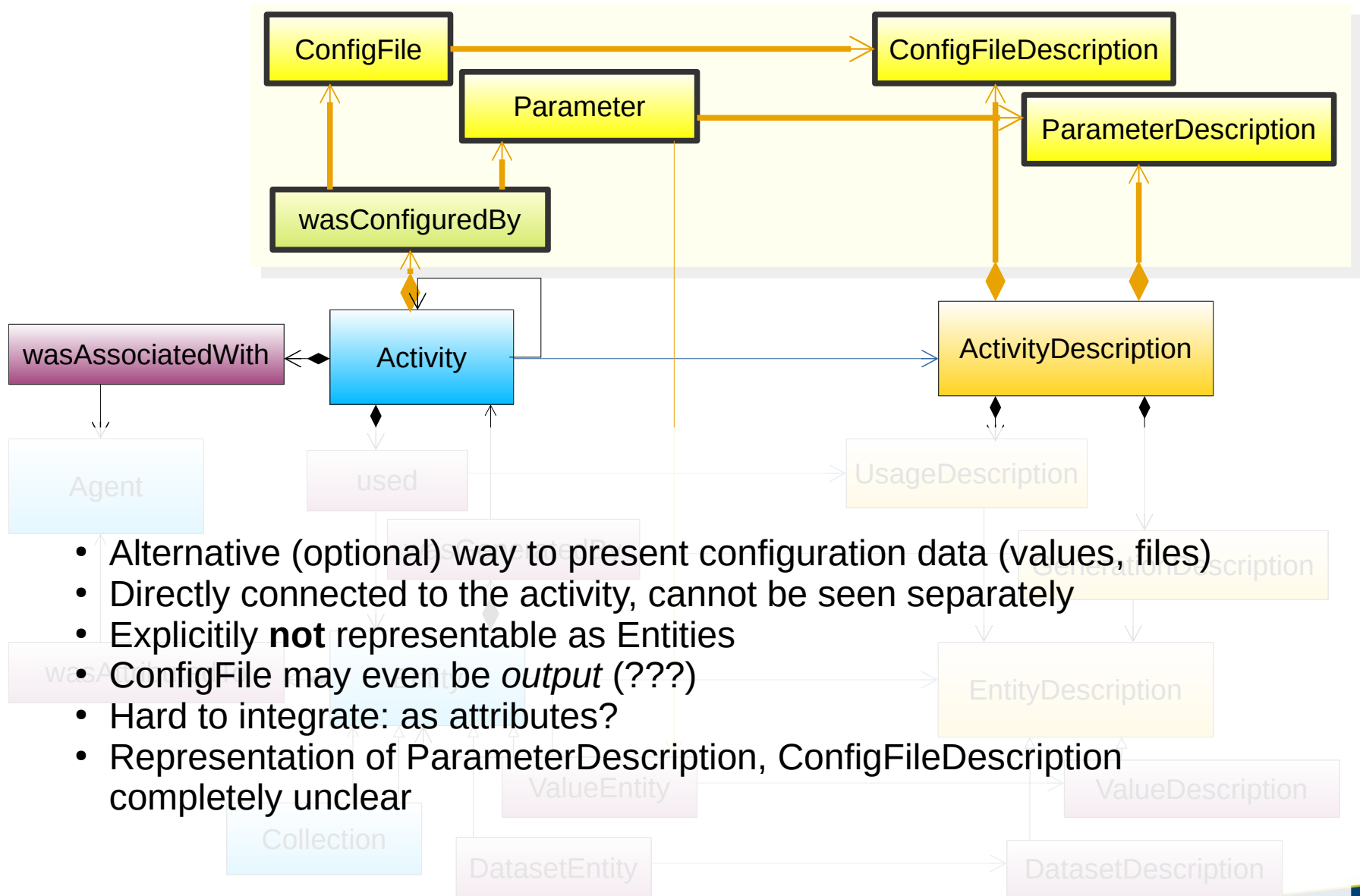
- ValueEntity, DatasetEntity
- ValueDescription, EntityDescription
- Collection, hadMember
- Unproblematic (prov:Type attribute of Entity)



ActivityConfiguration



ActivityConfiguration



- Alternative (optional) way to present configuration data (values, files)
- Directly connected to the activity, cannot be seen separately
- Explicitly **not** representable as Entities
- ConfigFile may even be *output* (???)
- Hard to integrate: as attributes?
- Representation of ParameterDescription, ConfigFileDescription completely unclear