



IVOA PROVENANCE DM CURRENT STATUS

Kristin Riebe, Anastasia Galkin , Ole Streicher, AIP

Mathieu Servillat, LUTH

François Bonnarel, Mireille Louys, CDS

Michèle Sanguillon, LUPM,

Markus Nullmeier, Ari , Heidelberg , Gavo

Laurent Michel, SVOM, Observatoire Strasbourg
and the IVOA Data Model Working Group

Handling Provenance representations



- We re-use results from W3C provenance
 - Concepts → part of IVOA DM
 - W3C Visualisation Tools
 - Southampton Provenance suite
 - Code and Libraries Python PROV
 - Serialisation Formats :
 - PROV-N , PROV-JSON, PROV-XML , ~~PROV-O~~

Provenance in the W3C[®]

- **4 recommendations (30/04/2013)**

PROV-DM: the PROV data model → *extensible with new attributes*

PROV-O: the PROV ontology

PROV-Constraint: Constraints of the PROV Data Model

PROV-N: a notation for provenance aimed at human consumption

- **and a number of non-prescriptive notes**

PROV-XML: an XML schema for the PROV data model

PROV-AQ: Provenance access and query

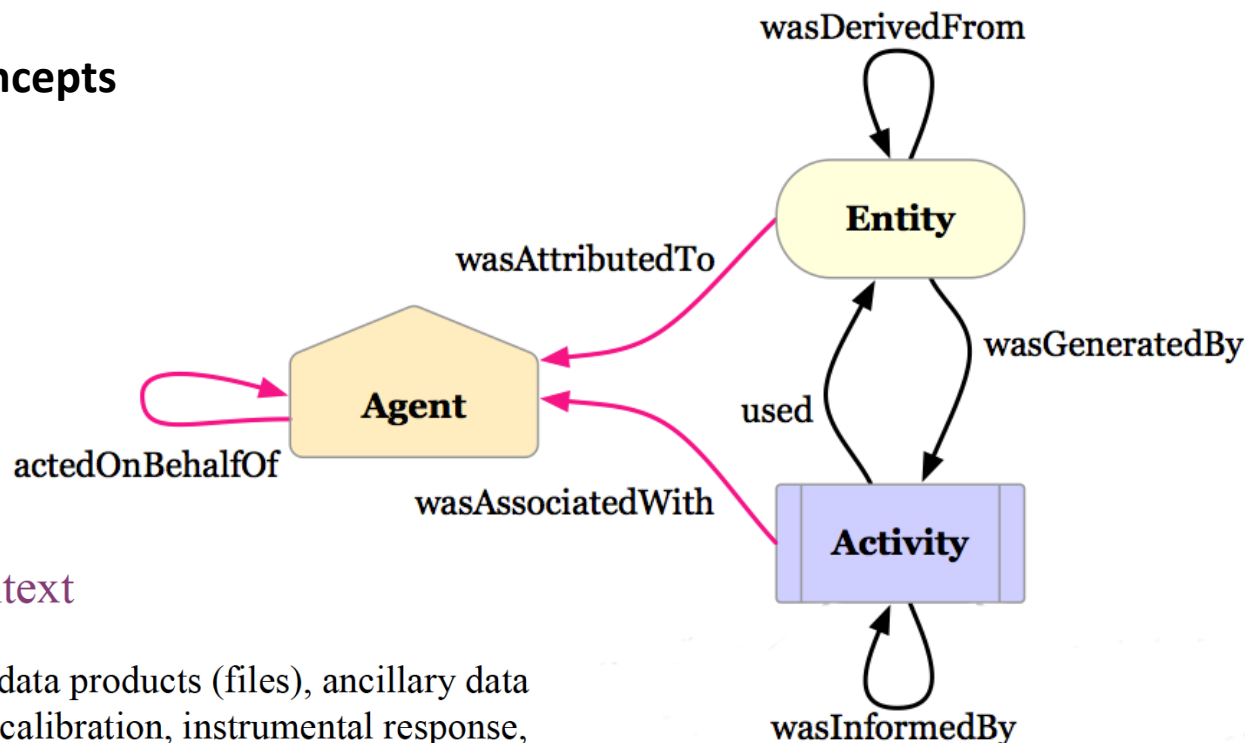
- **Tools to translate from one format to another & experience**

Southampton Provenance suite

<https://provenance.ecs.soton.ac.uk/>

Visualization

Core Concepts

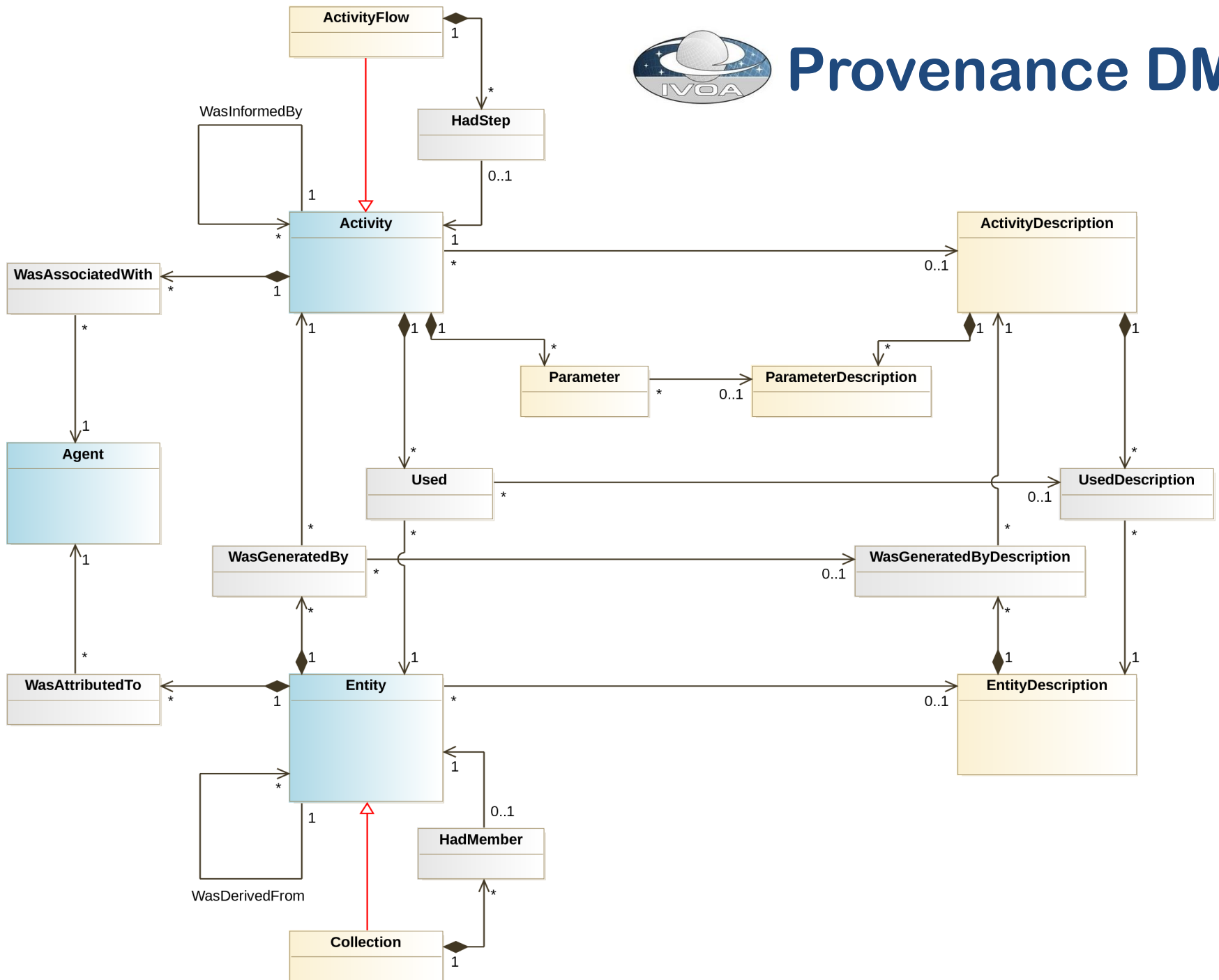


In our context

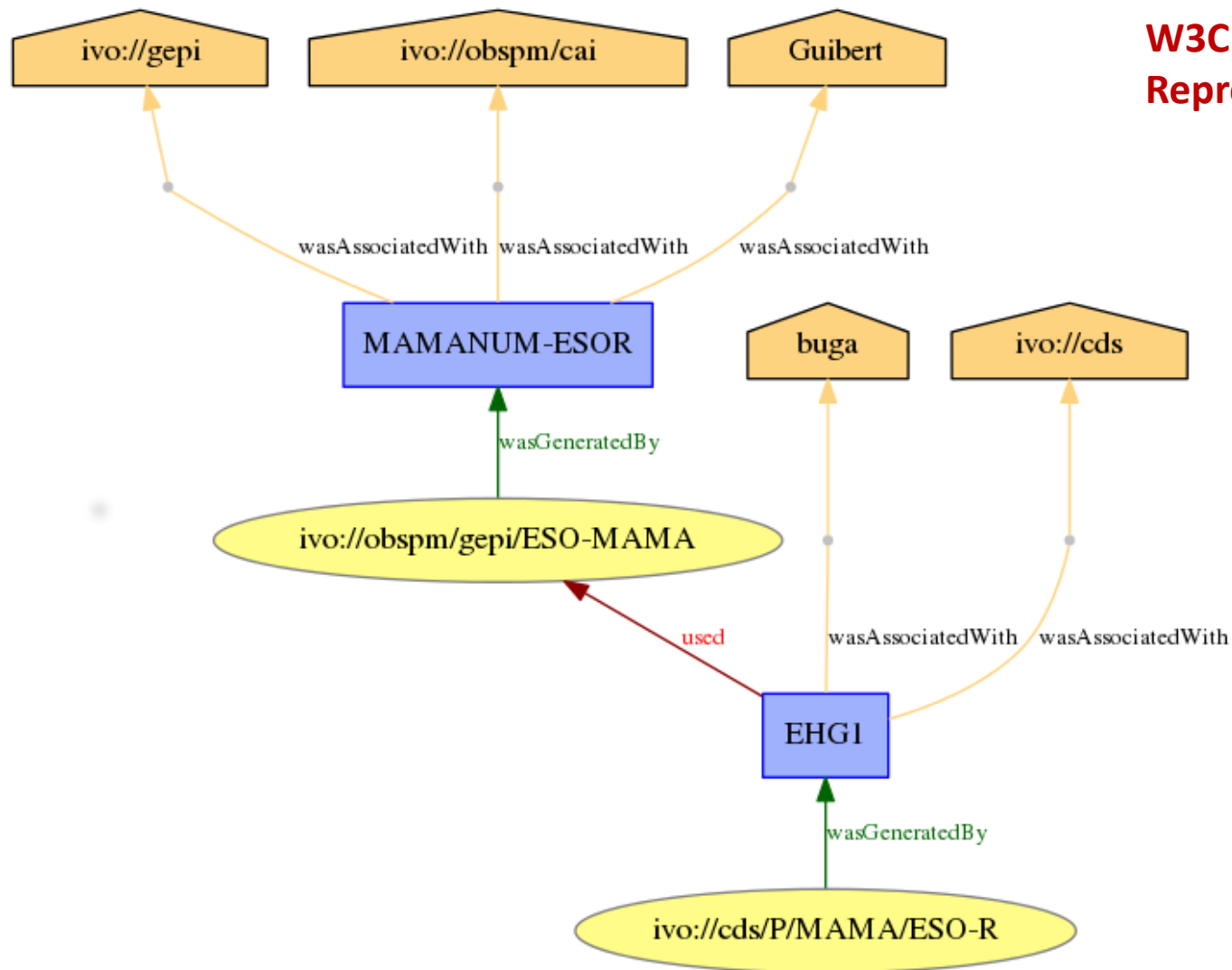
- Entity** ■ data products (files), ancillary data (calibration, instrumental response, etc.), processing parameter files
- Activity** ■ data acquisition, mosaicing, regriding, fusion, calibration, ..., transformation
- Agent** ■ Telescope astronomer, pipeline operator, principal investigator, etc.



Provenance DM



□ HiPS Generation use case



**W3C Graphical
Representation**

Example of the PROV-N format

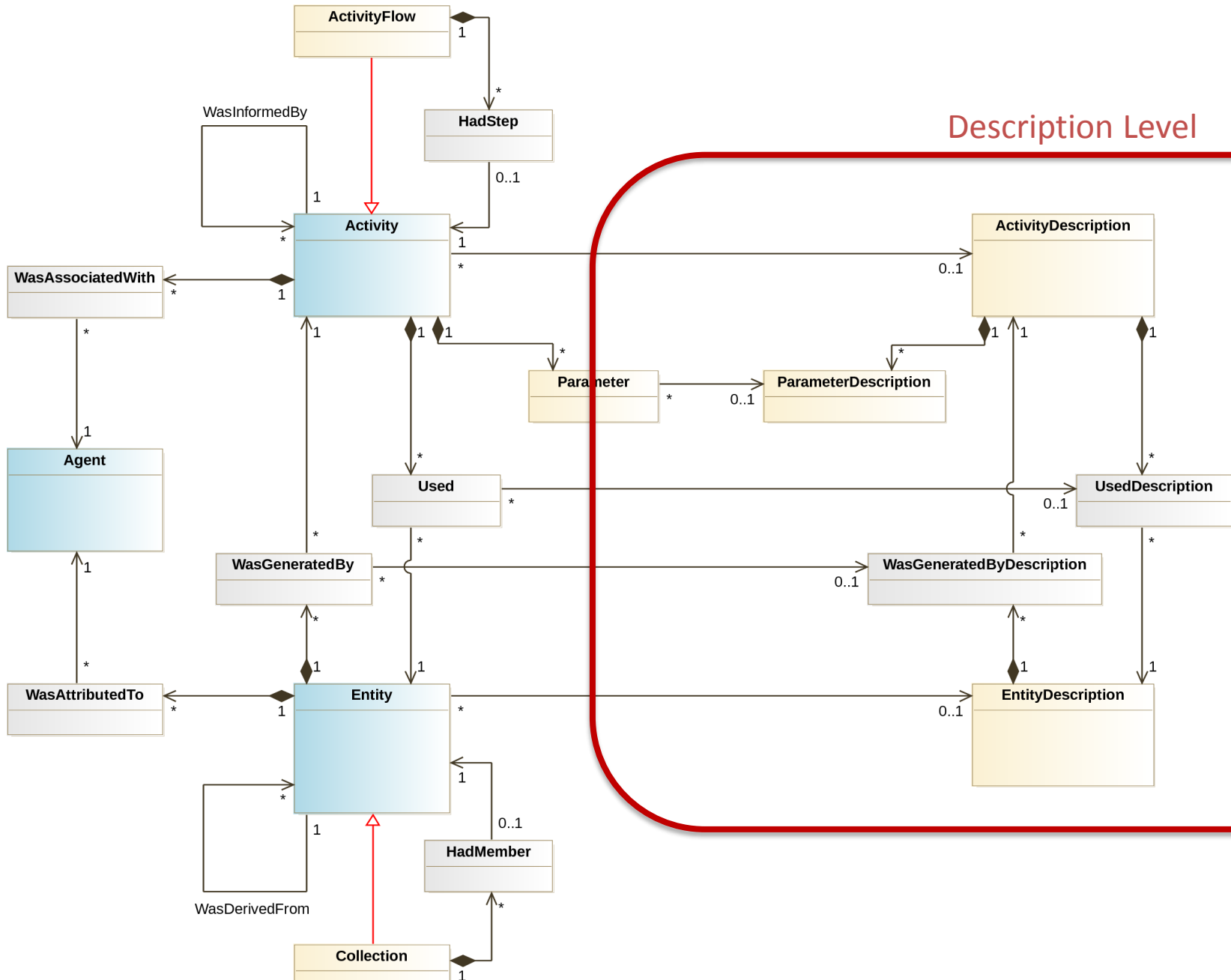
**Logical representation as tuple
in a relational DB**

```
activity(MAMANUM-1, date1, date2)
entity(ivo://SERC/Plate#SERC.J.444)
used(MAMANUM-1, ivo://SERC/Plate#SERC.J.444, -)
wasGeneratedBy(ivo://gepi/MAMA/Num#SERC.J-MAMA.444, MAMANUM-1, -)
entity(ivo://gepi/MAMA/Num#SERC.J-MAMA.444)
used(cds_cutoutj444, ivo://gepi/MAMA/Num#SERC.J-MAMA.444, -)
activity(cds_cutoutj444, -, -)
wasGeneratedBy(ivo://cds/P/MAMA/SERC#SERC.J-MAMA.444, cds_cutoutj444, -)
entity(ivo://cds/P/MAMA/SERC#SERC.J-MAMA.444)
used(AlaRGB6, ivo://cds/P/MAMA/SERC#SERC.J-MAMA.444, -)
activity(AlaRGB6, -, -)
entity(ivo://cds/P/MAMA/ESO#ESO.R-MAMA.444)
used(AlaRGB6, ivo://cds/P/MAMA/ESO#ESO.R-MAMA.444, -)
entity(ivo://cds/P/DSS2/SERC#SERC.I-DSS2.445)
used(AlaRGB6, ivo://cds/P/DSS2/SERC#SERC.I-DSS2.445, -)
wasGeneratedBy(ivo://cds/P/DSS2color#RGB_M83, AlaRGB6, -)
entity(ivo://cds/P/DSS2color#RGB_M83)
```


Core DM Transcription

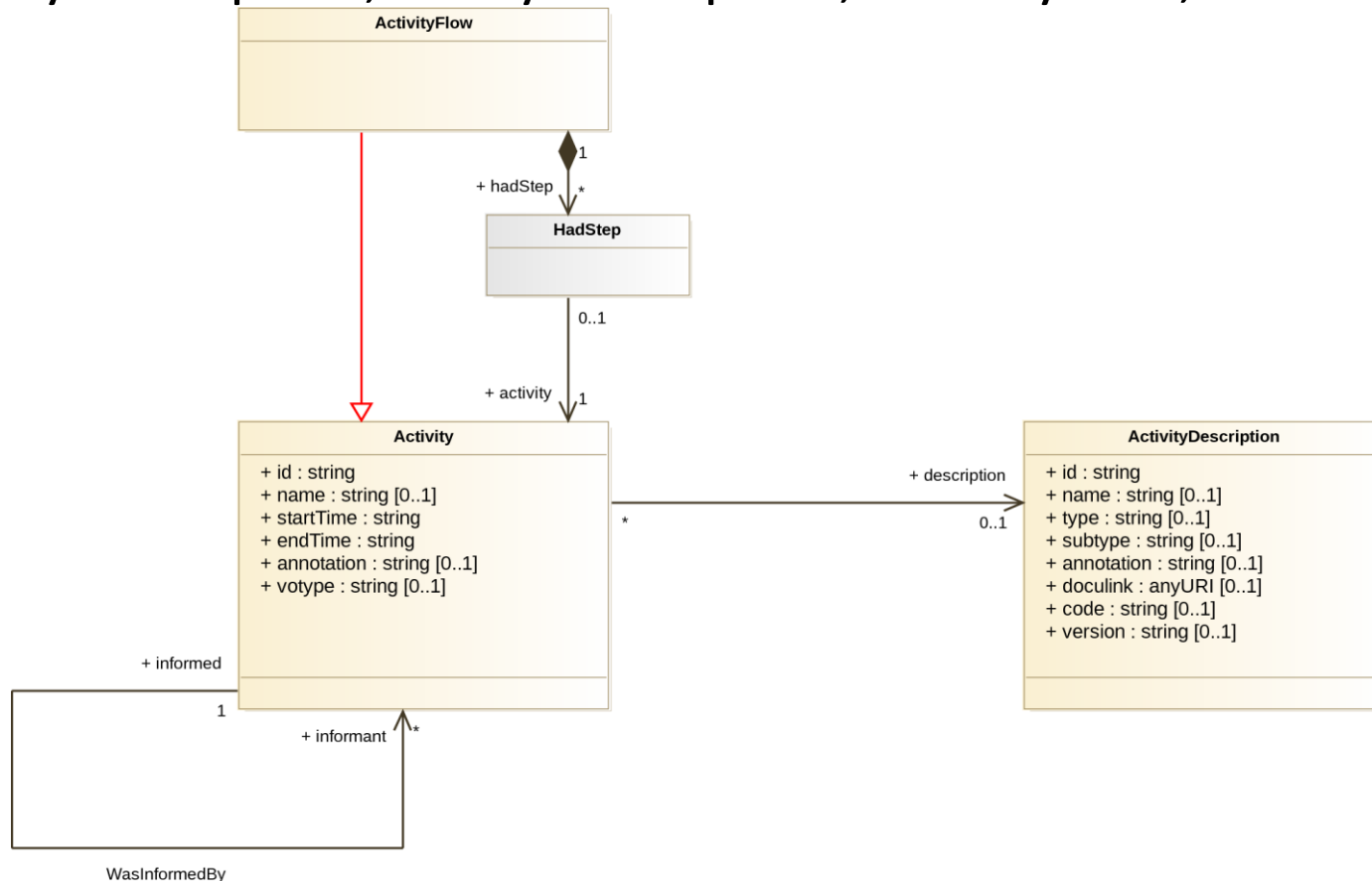
- In W3C
 - All classes represented in IVOA DM also exist in W3C
 - Easy to express in all PROV W3C PROV-XML, PROV-N, PROV-JSON
- in IVOA
 - Easy to express in IVOA multiple tables (VOTable)
 - TAP compatible

IVOA Provenance DM



IVOA Classes for Description

- Enrich the W3C classes by adding new classes
ActivityDescription, EntityDescription, ActivityFlow, etc



In the VO ecosystem

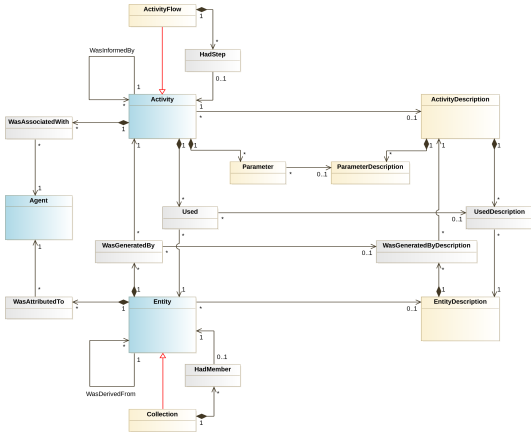
The screenshot displays a complex VO ecosystem interface with several overlapping windows:

- Agents:** A table listing various astronomical observatories and their associated organizations.
- Activity description:** A table showing activity details such as label, type, subtype, and URL.
- Topcat:** A window showing a table list and current table properties for a specific dataset.
- Entity:** A detailed table browser for a specific dataset, showing columns for ident, label, type, and annotation.
- Activity:** A table browser for a specific activity, showing columns for ident, label, start, stop, and annotation.
- Parameter:** A table browser for a specific parameter, showing columns for id, value, unit, and ucd.
- WasGeneratedBy:** A table browser for a specific activity, showing columns for head, tail, and role.
- Used:** A table browser for a specific activity, showing columns for head, tail, and role.
- MaxAssociatedWith:** A table browser for a specific activity, showing columns for head, tail, and role.

Specific classes Transcription

- in IVOA ecosystem
 - Use one table for each defined class
 - fully extensible
- In W3C
 - No W3C structure corresponding to added IVOA classes
 - Need some transcription to the W3C existing constructs
 - **Refactoring of the IVOA PROV DM for a PROV W3C view**

IVOA DM



All classes defined in the specification

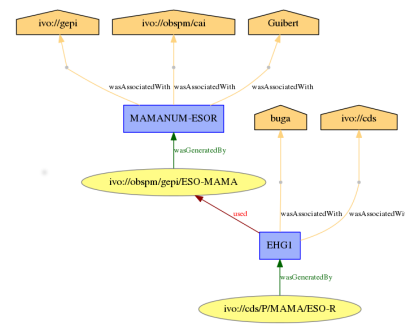
Translation Projection

Translation Projection



PROV-JSON
PROV-N
PROV-XML

W3C Tools
Provenance suite, Prov-Python



Prov-DAL

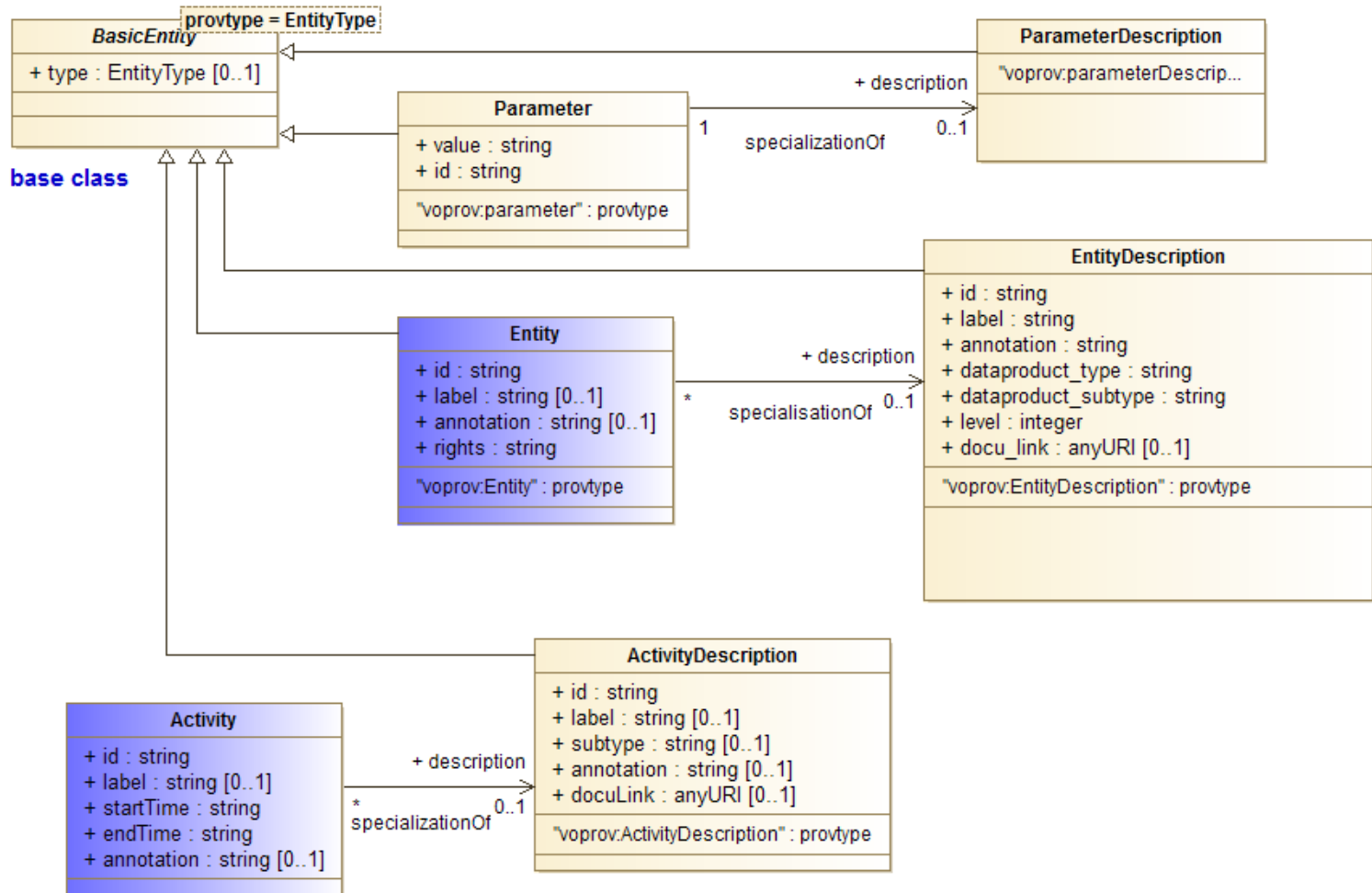
Prov-TAP

TopCat
TapHandle

VOTable
JSON
FITS

Entity re-use for description

W3C PROV
View



Current status of IVOA Provenance DM

- A Core Provenance Model proposed by the IVOA data model working group at ivoa.net/Documents
See [WD-ProvenanceDM-1.0-20170921](http://www.ivoa.net/Documents/WD-ProvenanceDM-1.0-20170921)
Discussed last fall / little feedback on ivoa lists
- Will be split in lighter more focused docs
 - Data Model proposed REC
 - Implementation Note
 - Prov-TAP access protocol for Prov by TAP (DAL)
 - Prov-DAL access protocol for Prov by DAL (DAL)

IVOA PROV Implementations

Based on four use-cases

- [CTA](#) (M. Servillat)
- [RAVE – the Radial Velocity Experiment](#)
- [POLLUX](#) (synthetic stellar spectra service)
- [SVOM](#) gamma ray burst /transients
- Prototype TAP-based API for images in an archive (@CDS)
- [MUSE](#) Example (see Ole's slides)

Various usage of the model

- Lessons learnt from implementing prototypes

Project/ implemented features	Protocol	Activity Flow Multi-level Activity	Activity /Entity description / Parameters	Serialisation Formats
RAVE	Prov-DAL	Yes	Yes	PROV-JSON , PROV-N
CTA	TAP/ UWS	Yes for future	Yes	VOTable, PROV-JSON , JSON
SVOM	Prov-DAL?	?	yes	PROV-JSON , JSON in Fits
POLLUX	Adhoc then Prov-DAL	No	Yes	All
MUSE	AstroWise	?	Yes as entities?	?

- Various profiles to propose : Workflow view, data flow view, credits/rewards view , etc ..

Provenance WP management



Provenance days : meeting site

<http://wiki.ivoa.net/twiki/bin/view/IVOA/ObservationProvenanceDataModel>



Reports and discussions at DM sessions during the IVOA Interoperability meetings



Connections to interdisciplinary projects
See WG for « Provenance Patterns »

Broader connections



Prov Patterns WG

Prov Patterns WG Collaboration Tools

- <https://www.rd-alliance.org/group/provenance-patterns-wg/wiki/prov-patterns-wg-collaboration-tools>
- Gathering Provenance use cases across various disciplines → A Use-case DB
- A collection of Provenance patterns to answer elementary use-cases → building blocks solutions
 - Prov-O (ontologies) , RDF/ Triple store oriented
- Contributions from the IVOA very welcome

Topics for discussions

- Parameter representations
 - Views in IVOA
 - Views in W3C
- The various profiles the IVOA DM can offer
 - Workflow : Activity focused
 - Data flow / archive : Dataset focused
 - Credits/responsibility views
- Protocols for accessing Provenance
 - Prov-DAL
 - PROV-TAP
- Others : please bring up yours ...