

SimDM and the Registry

Hendrik Heint, INAF-OATS
Astro-CC TechForum, Trieste

October 9, 2025

Agenda

What's SimDM ?

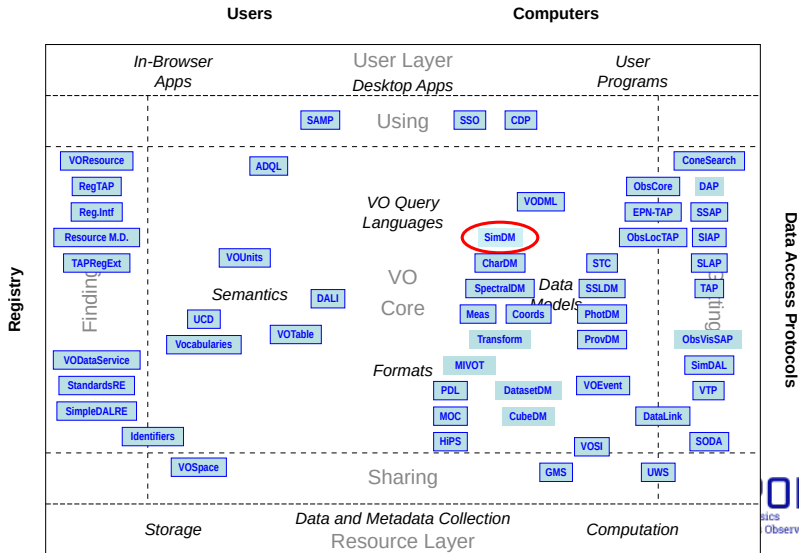
Is this science or can it be thrown out?

Vocabularies

Preparing for AI

Quo vadis SimDM ?

SimDM in the VO architecture



Graphical view on SimDM



Going through SimDM, one sees it's overwhelming and on the service side can cover a lot of questions in variable granularity towards making a selection to the data (e.g. for specific physics down to single parameters). The focus of SimDM seems to be to describe the process of producing a dataproduct.

A point I am missing a bit is: what kind of question is a user supposed to ask to the VO-Registry to actually find a Service with SimDM ?

What would we need to make this meaningful ?

Lost of words

Describing Physics (and Applied Physics) demands a vocabulary. Unfortunately the links within the documents are broken (see presentation by Baptiste Cecconi for possible solution).

Use our own words ?

How do we solve this ? Will we define an IVOA vocabulary for this, which will have to be maintained ? Do we trust a different third party to maintain theirs but risk to run into the same trap ?

Anyway: for SimDM as it is now, we will at least need an erratum with a corrected link to the Vocabularies.

Is high granularity a curse...

Describing an experiment down to the level of single parameters (of which easily dozens can be set) and their dependencies in regard of the simulated physics and the applied algorithms from the protocols is a tedious task. This may be one of the reasons why SimDB did not take off as hoped.

... or a blessing?

In Preparation of AI, this granularity would enable a machine to learn more details about the simulated physics and support more detailed analysis of simulation results. So skipping totally would seem like an overreaction.

Short/middle term goals

- ▶ correct the Vocabularies with an erratum
- ▶ find a reliable way to maintain such vocabularies (Baptiste Cecconi already working on this).
- ▶ find enough examples to test of what works for SimDB in regard of the VO-registry (volunteers ?)
- ▶ reviving the Theory IG for discussion (volunteers ?)

Long term goals

Having a few Simulation Data Services running exposing dataproducts in a standardized way to (at least) enable good science, or (at best) have datasets available as training sets for AI.

References

- ▶ SimDM standard
- ▶ Graphic view of SimDM
- ▶ Browseable specification of SimDM