SimDAL implementation for Pollux

16 May 2019 - Paris

O. Tortosa - M. Sanguillon - A. Palacios

Internship’s subject

- ➔ 2-year university degree
- ➔ 10 weeks
- ➔ SimDAL and SimDM for Pollux

Pollux

- synthetic stellar spectra database (only metadata)
- currently using SSAP and ProvSAP protocols
- raw data access thanks to:
  - pollux.oreme.org
  - all tools that implement SSAP (for example: Cassis)

Goals

Why for Pollux?
To acknowledge Pollux data as simulations in addition to also be spectra

- SimDM:
  - describing all simulations in the same way and in great details
- SimDAL:
  - search for an object without knowing the simulation a priori

Actual state of implementation of SimDM/SimDAL

Implemented in only 2 laboratories:

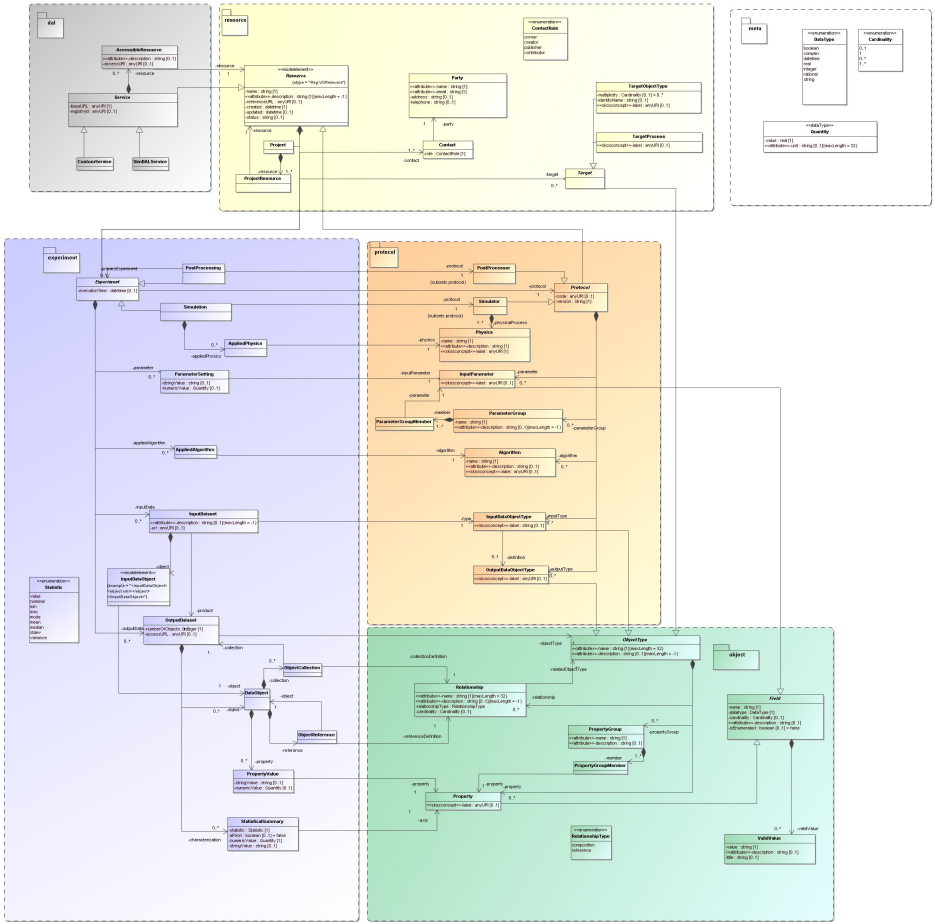
- Paris Observatory (SimDAL)
- Spanish Virtual Observatory (SimDAL)

Not really used at the moment because there is no client yet

Only previous versions of simDM has been implemented

SimDM

Keep the current structure of the Pollux database and not implement the complex SimDM



SimDAL architecture: 3 parts

Repository:

- ➔ Allows the search for an object through keywords
- ➔ Acts as yellow pages and redirects to individual databases

Search:

- ➔ Allows to choose a project and see the available parameters
- ➔ Researchers can select on parameters

DataAccess:

- ➔ Allows to see statistics of a dataset and retrive a simulation
- ➔ Researchers can select on dataset characterics to choose a part of the simulation or raw data

Repository: SimDM resources

Projects:

Describe the projects available in the database

Are used for the repository search thanks to key words

Protocols:

Describe algorithms, physical processes and parameters taken into account to generate the simulation

Experiments:

One for each simulation

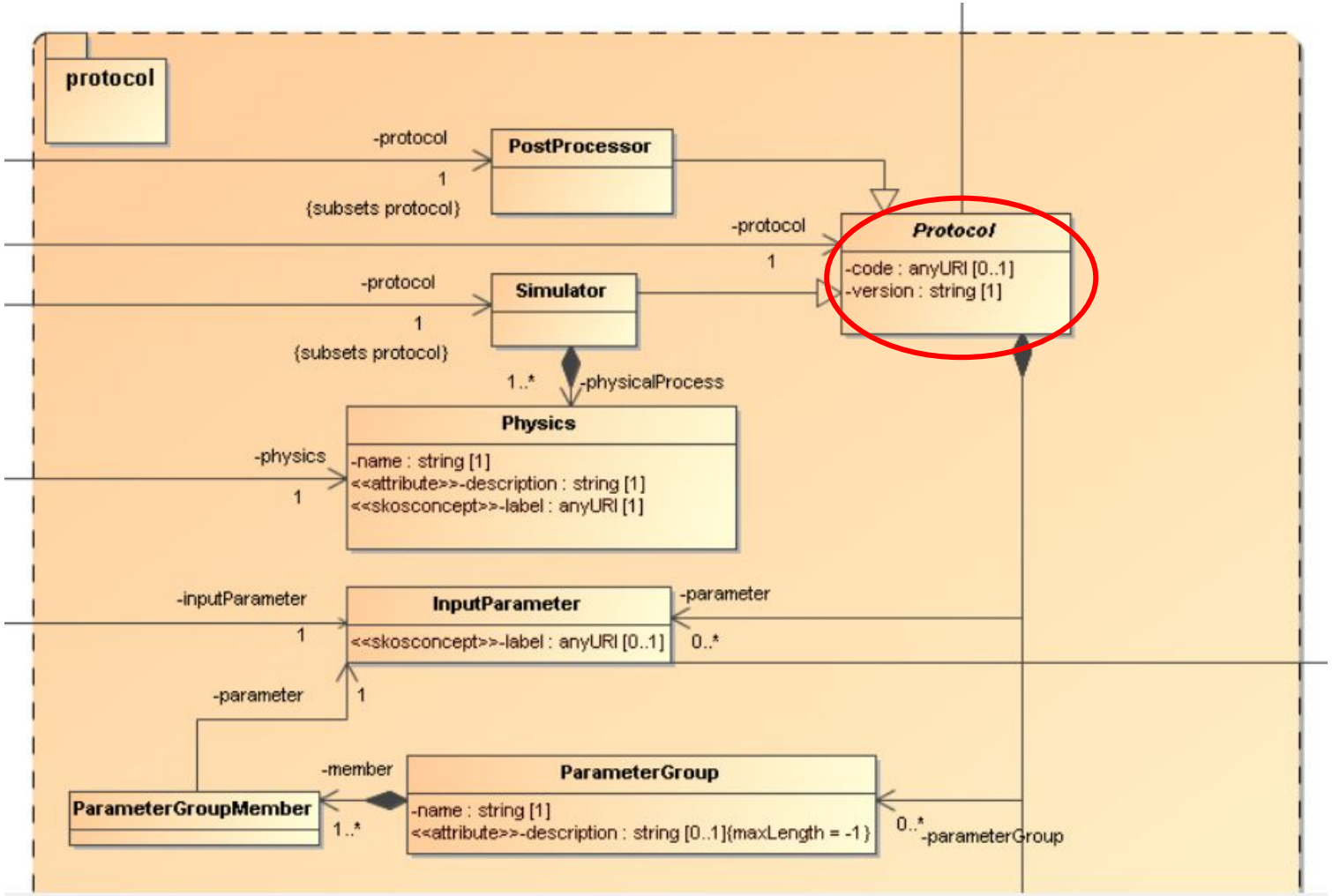
Define the values corresponding to the protocol’s attributes

Choices made

Use of Paris Repository

Regarding SimDM resources:

- only Projects and Protocols for referencing
- 1 project = 1 collection
- 1 protocol contains 2 algorithms (atmospheric model and spectral simulation)


protocol
-protocol
1
{subsets protocol}
PostProcessor
-protocol
1
Protocol
-code : anyURI [0..1]
-version : string [1]
-protocol
1
{subsets protocol}
Simulator
1..*
-physicalProcess
Physics
-name : string [1]
<<attribute>>-description : string [1]
<<skosconcept>>-label : anyURI [1]
-physics
1
-inputParameter
1
InputParameter
<<skosconcept>>-label : anyURI [0..1]
-parameter
0..*
-parameter
1
-member
ParameterGroupMember
1..*
ParameterGroup
-name : string [1]
<<attribute>>-description : string [0..1]{maxLength = -1}
0..*
-parameterGroup

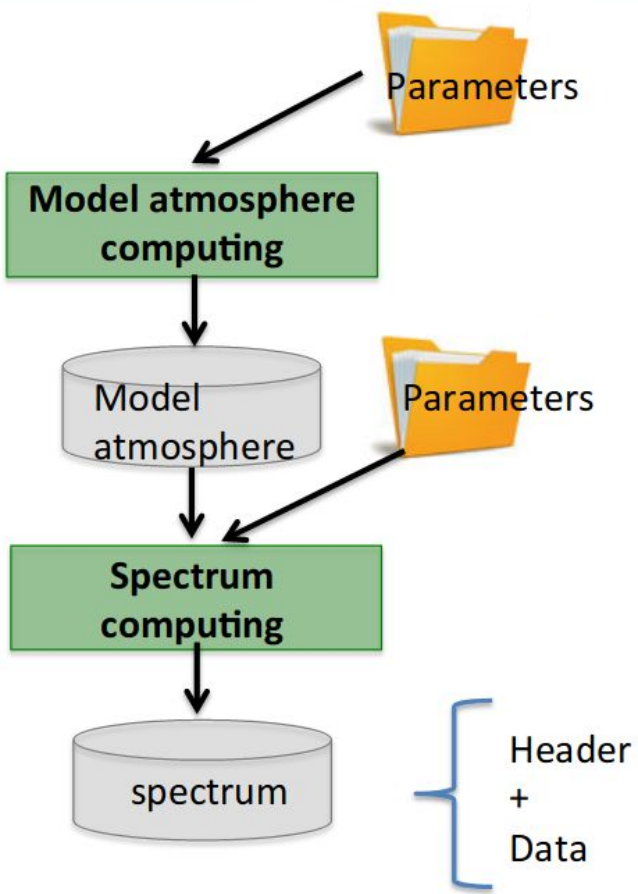
Choices made

Use of Paris Repository

Regarding SimDM resources:

- only Projects and Protocols for referencing
- 1 project = 1 collection
- 1 protocol contains 2 algorithms (atmospheric model and spectral simulation)

=> 1 code? ideally 2 codes? is the output for the first code necessary?


Parameters
Model atmosphere computing
Model atmosphere
Parameters
Spectrum computing
spectrum
Header
+
Data

Project Management

Guidelines: Being able to look for a simulation and to retrieve it

Adding finer selection criteria along the way

- **Repository:**
  - **Projects**
  - **Protocols**
- **API search:**
  - **views**
  - **schéma**
  - **cutout**
- **API dataAccess:**
  - **dataset**
  - **schéma**
  - **cutout**
  - **rawdata**

Step 1
Criteria = collection

Step 2
Criteria = parameters

Step 3
Criteria = dataset statistics

Current state of the project

Still 4 weeks to go

Repository: SimDM resources

Projects:

Not done yet

Protocols:

Done with 2 codes in one resource

Vocabulary: non Local Thermodynamic Equilibrium? or nothing?

API search

Views:

Possibility to see all collections

Schema:

Possibility to see each schema

Vocabulary : Plane-parallel Geometry? Spherical Geometry?

Cutout:

Cutout without JSON and only one attribute as a criterion

API dataAccess

Dataset:

Currently not available after cutout

Links in Views for all simulations in a collection

Schema:

Not implemented yet

Cutout:

Not implemented yet / will only be possible on the wavelength

Raw data:

Available

Issues summary

- IVOA vocabulary missing: nonLTE and Geometry (spherical or plane-parallel)
- 2 algorithms in 1 Protocol
- overly difficult to implement for simple simulations
- worth it for Pollux?
- lack of tools: VOTables are not user friendly

Special thanks to David Languignon for his help and guidance

Thank you for your attention
Any questions? Any advices?