Theory I.G.

SimDAL

Franck Le Petit, David Languignon
Gerard Lemson, Hervé Wozniak

Hawaii - InterOp - 27 september 2013

Scientific use cases

Find a synthetic spectrum of O 9 star

Find a list of properties of halos (masses, positions, velocities) in a LCDM model at z about 3.

Extract a piece of cosmological simulations around a halo

Preparation of observations :
Find estimations of line intensities of CO 1-0, 2-1 and 15-14 in a cloud of density of 10^5 cm^{-3} and different illuminations.

Interpretation of observations
What are the physical conditions in interstellar clouds that reproduce :

- 1E20 < N(H2) < 3E20 cm-2
- N(CO) > 2E14
- N(CH+) > 1E14

Simulation Data Model

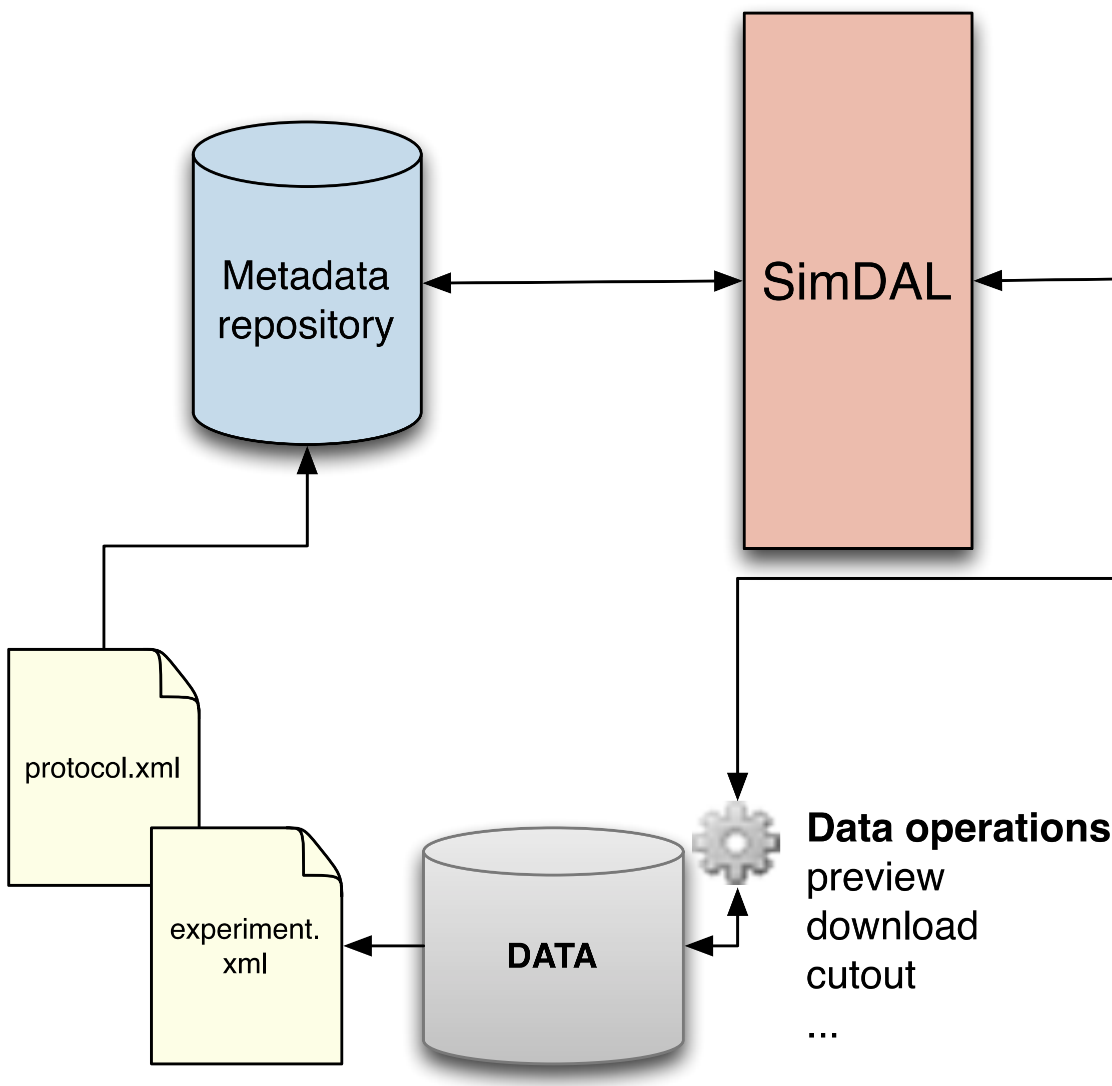
SimDM : precise description of codes / simulations
- Input Parameters
- Properties
- Statistics
- ...

Descriptions of entities :
- SKOS concepts
- text fields (names / descriptions)

Definition - Properties
Quantity computed at each position / for each object in a simulation.
Example : velocity, position, mass
They can be very numerous :
Example : line intensities CO 1-0, 2-1, 3-2, ...

Specificities of Theory :
- **heterogeneous multi-dimensional data / meta-data**
- (eventually) **voluminous data** (Ex: cosmology > To)
- (eventually) **lots of meta-data** (Ex: line intensities / PDR > 150 000 columns)

Architecture for Theory services in the VO



Discovery is done in 2 steps :
- Services discovery
- Datasets discovery

Then we retrieve data :
- Download
- Extraction of data / cutout
- ...

Service discovery

Examples :

- *Search theoretical services about* «stellar atmospheres»
- N-body *simulations modeling* halos

Step ① : Search using

- tags
- keywords
- concepts (SKOS)

→ ID of resources
+ matching fields

Step ② : Get meta-data

- ID of the resource

→ Metadata of the resource
Format :
protocol.xml
experiment.xml
project.xml
service.xml

Service discovery

Comparison to **Registry Interface**

Registry Interface	
• Search	✓
• KeywordSearch	✓
• XQuerySearch	✗
Pages for large number of results	✓

- For theory, a single resource can produce a large file
  Risk of slow communication client / server

Example :
PDR services (150 000 properties because of line intensities) - Resource description : 30 Mb

➡ **Need pagination for resources**

SimDB Interface
- Search
- KeywordSearch
- Pages for large number of resources
- Full / Light version of resource descriptions

Full description : everything
Light description : no collections
- Collections can be asked later

Implementations

Discovery of services

- A few functions as Registry Interface
- Get meta-data Full or Light versions

Implementation

- Free back-end implementations (RDBMS, object oriented DB, ...)
- **REST compatible**

`/protocols`	List of protocols
`/protocols/id`	meta-data of a protocol
`/protocols/light/id`	Light meta-data of a protocol
`...`	...

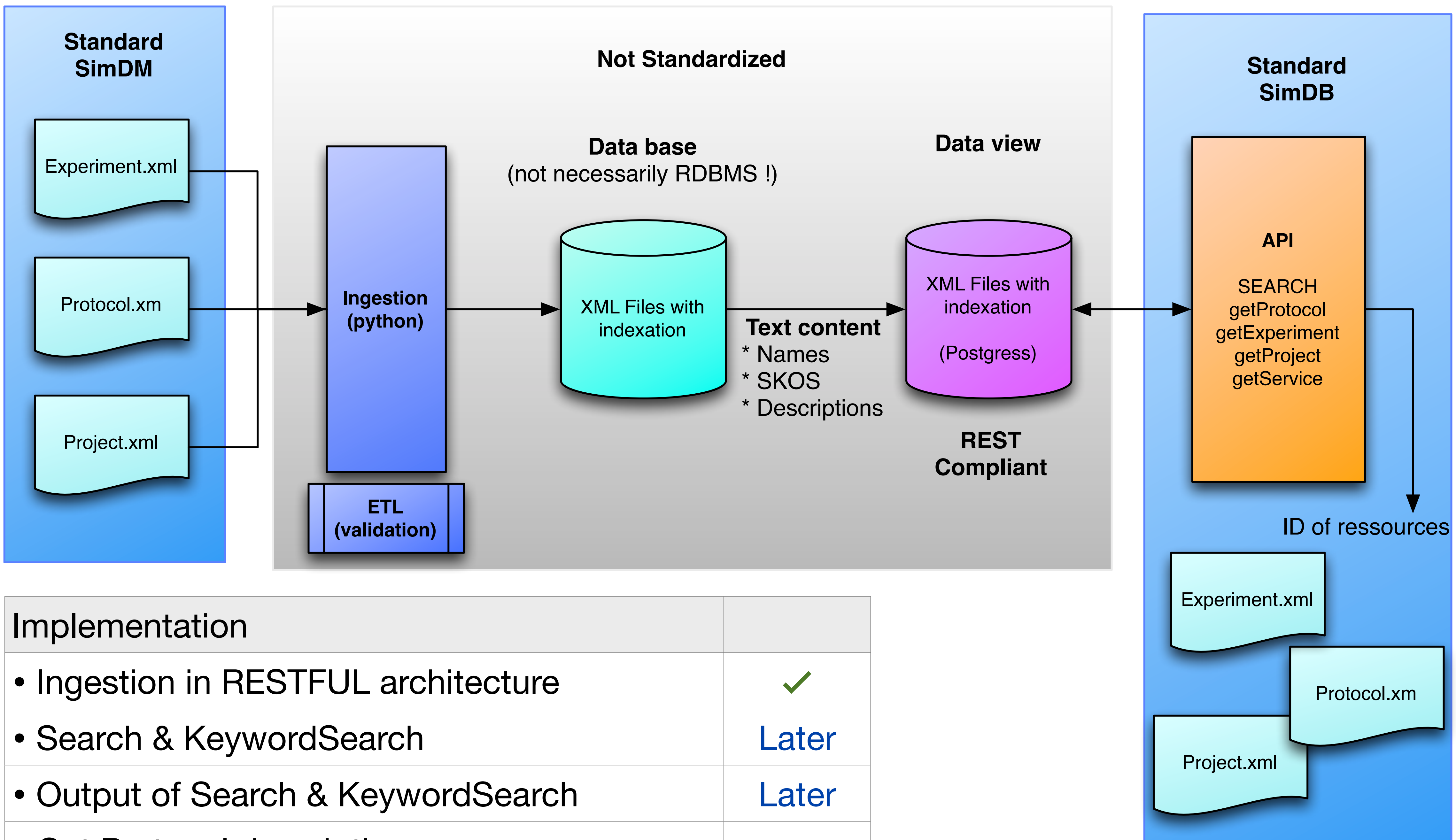
Data exchanged format

SimDM serialization

- protocol.xml
- experiment.xml
- service.xml

Implementations of Service Discovery

Example of implementation in VO-Paris



Implementation	
• Ingestion in RESTFUL architecture	✓
• Search & KeywordSearch	Later
• Output of Search & KeywordSearch	Later
• Get Protocol description	✓
• Get Experiment description	✓
• ...	
• Output serialization	✓

Discovery of datasets

Goal : Discover a dataset

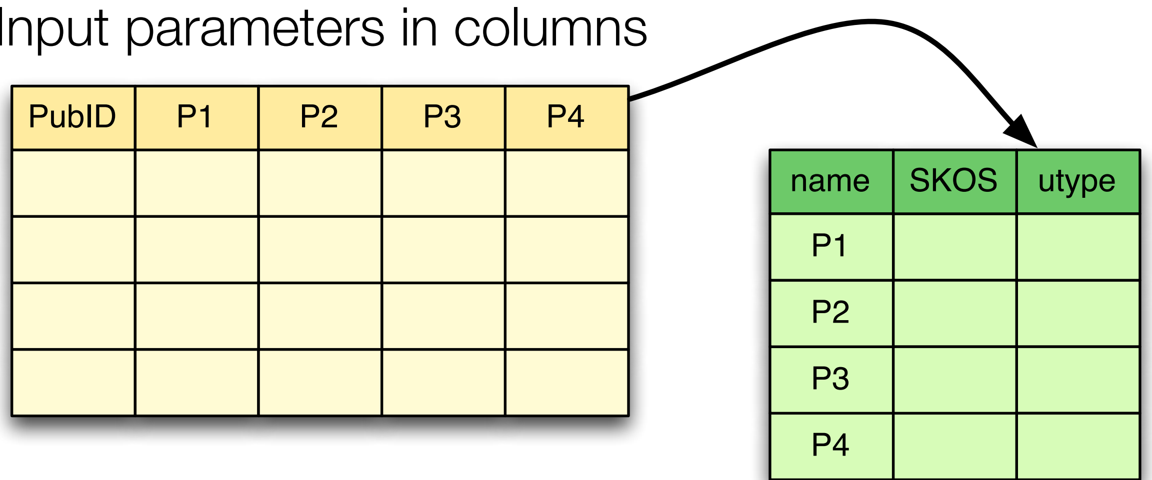
At this step of the discovery, numerical values are used

Example : Find stellar spectrum of **0 9** type
Find halos with **mass** > **X solar mass**

Initial idea : **SimTAP**

TAP queries on SimDM are complex
SimTAP : pivoted tables to simplify queries

Input parameters in columns



TAP schema

Implementations :

✓ OK for simulations with few quantities (InputParameters, Properties)

✗ But NOT for services with lots of Properties / InputParameters

Example :
Services publishing computed line intensities
PDR : more then 150 000 columns

Tables in RDBMS / SQL too big for efficiency

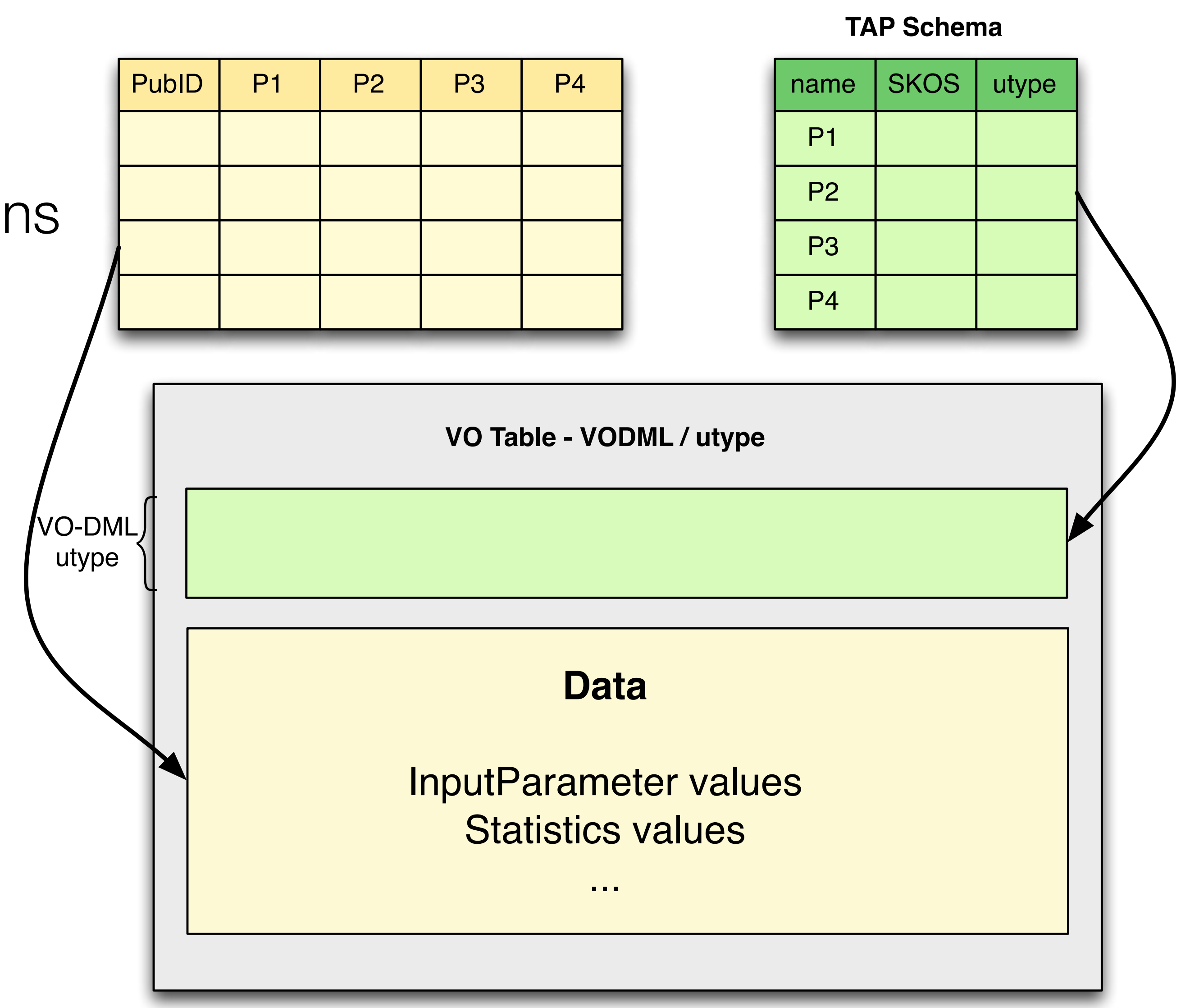
➡ **Need solution independent of data management system**

Discovery of simulations

SimDAL

Philosophy

- Publishers store meta-data as they wish
  - Meta-data considered to be stored in virtual tables / VO-Tables
  - be able to produce VO-Tables :
    - files as described in VO-DML - utype
- Interface :Few operations to discover simulations
  - Simple REST service end points
  - Simple Query Language



Query Language

What do we need ?

votable

"select ... from ... where ..."

field1, field2, ...

constraint1, constraint2, ...

field operator value

- Need only **select**, **from** and **where**

Datalink

Description of possible operations on datasets

```
▼<votable xmlns="http://www.ivoa.net/xml/VOTable/v1.2" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="1.2">
 ▼<resource>
  ▼<table>
     <field ucd="meta.id" utype="datalink:Datalink.uri" name="uri" arraysize="*" datatype="char" xtype="w3c:URI"/>
     <field utype="datalink:Datalink.accessURL" name="accessURL" arraysize="*" datatype="char" xtype="w3c:URL"/>
     <field utype="datalink:Datalink.semantics" name="semantics" arraysize="*" datatype="char" xtype="w3c:URI"/>
     <field ucd="meta.id" utype="simdal:Artefact.productType" name="productType" arraysize="*" datatype="char"/>
     <field ucd="meta.id" utype="simdal:Artefact.contentType" name="contentType" arraysize="*" datatype="char"/>
     <field ucd="meta.id" utype="simdal:Artefact.contentLength" name="contentLength" arraysize="*" datatype="char"/>
   ▼<data>
    ▼<tabledata>
     ▼<tr>
      ▼<td>
         http://localhost:3000/datalink/cutout?uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
      ▼<td>
         http://localhost:3000/api/obspm/preview/hdf5?uri=http://localhost:3000/datalink/cutout?
         uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
       <td>ivo://ivoa.net/std/DataLink/v1.0#DOWNLOAD</td>
       <td>science</td>
       <td>application/hdf5</td>
       <td/>
     </tr>
     ▼<tr>
      ▼<td>
         http://localhost:3000/datalink/cutout?uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
      ▼<td>
         http://localhost:3000/api/obspm/preview/exp?uri=http://localhost:3000/datalink/cutout?
         uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
       <td>ivo://ivoa.net/std/DataLink/v1.0#PREVIEW</td>
       <td>science</td>
       <td>image/png</td>
       <td/>
     </tr>
     ▼<tr>
      ▼<td>
         http://localhost:3000/datalink/cutout?uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
      ▼<td>
         http://localhost:3000/api/obspm/preview/proj?uri=http://localhost:3000/datalink/cutout?
         uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
       <td>ivo://ivoa.net/std/DataLink/v1.0#PREVIEW</td>
       <td>science</td>
       <td>image/png</td>
       <td/>
     </tr>
     ▼<tr>
      ▼<td>
         http://localhost:3000/datalink/cutout?uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
      ▼<td>
         http://localhost:3000/datalink/cutout?uri=experiments&cutout=pubid=tdiff_n1e1r5m1r5m1z1e1a1p3_20_2012122310218
       </td>
       <td>ivo://ivoa.net/std/CutoutService/v1.0</td>
       <td>science</td>
       <td>application/json</td>
       <td/>
     </tr>
    </tabledata>
   </data>
  </table>
 </resource>
</votable>
```


Results of a discovery

URI of dataset

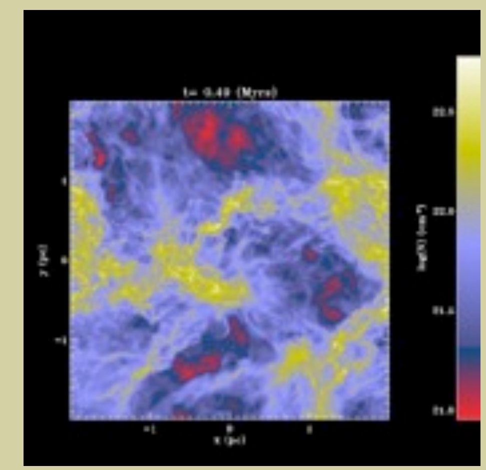
URI Download

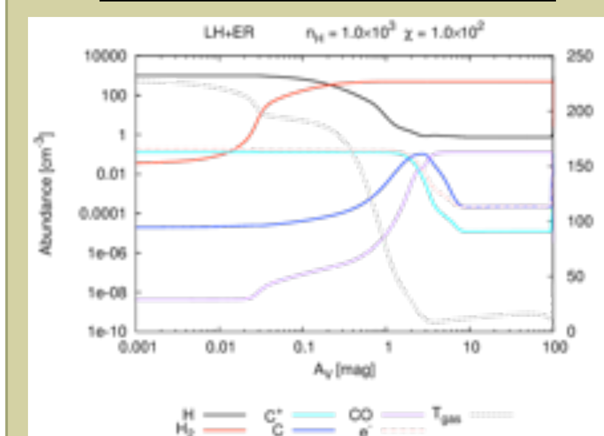
URI Preview Project

URI Preview experiment

URI Cutout

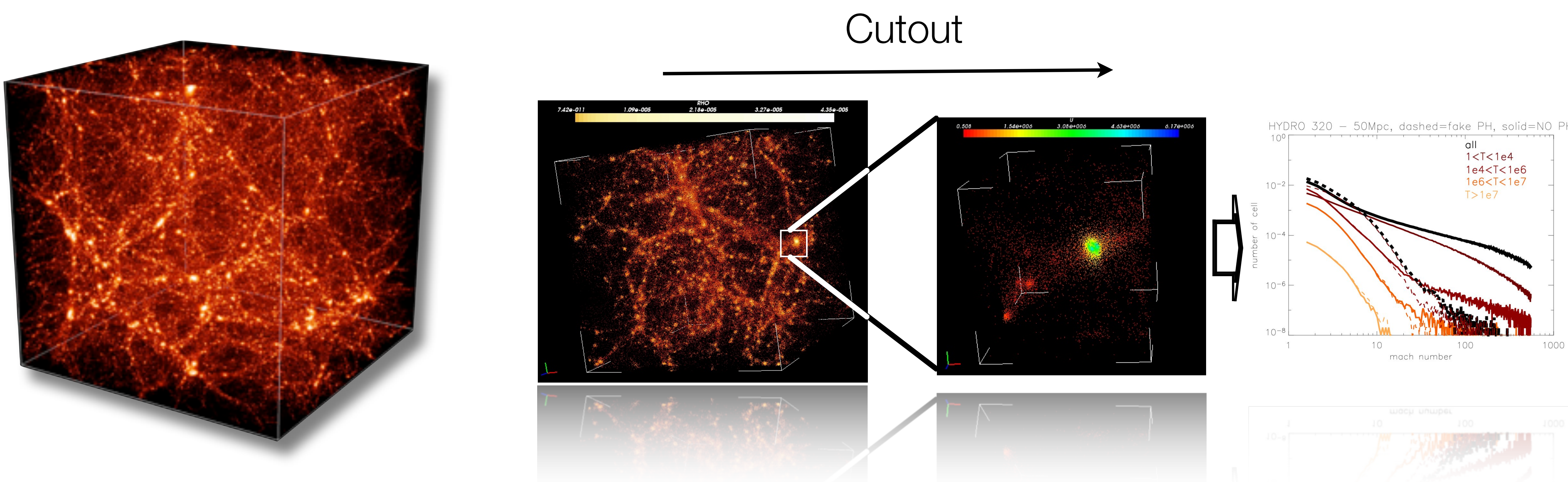
Preview





Cutout

Scientific use cases: ① Extraction in a data(hyper)cube



Large volumes of data
Need to extract a piece of the simulation

Simulations can be : 1D, 2D, 3D, other ?
We may wish to extract : cubes, spheres
Need projections in specific directions

Cutout

Scientific use cases : ② Catalogs

Example : radiative transfer
I observed CO 2-1 line intensities between values A and B erg cm^{-2} sr^{-1}
what are expected H_2O lines intensities ?

ID	CO 1-0	CO 2-1	CO 3-2	CO 4-3	...	H_2O 111-000	H_2O 110-101	H_2O 202-111	H_2O 212-101	...
1										
2										
3										
4										
5										
...										

Note :

Cutout can be done on many properties.

Description of cutout service may be big to list all possible quantities

Cutout

Requirements :

- Spatial coordinates are required
  - parallelepipedic in x, y, z
  - spheres : center + radius
  - 1D, 2D, 3D
- BUT ALSO non spatial coordinates
  - lambda (theoretical spectra)
  - x, y, lambda
  - time (continuous)
  - any quantities

- Need a generic cutout
- Need to specify properties to extract

Questions :

- Are all use cases taken into account in cutout specifications ?
- How to specify quantities to extract ?
- Where is the description of cutout service ?
  If the list of properties has to be provided in DataLink document => several Mo files

Exchanged files

Several files have to be shared between Client & Server

- Discovery phase : Service & datasets
  Files containing description of protocol, experiment, service, ...
  Proposition : protocol.xml, experiment.xml, ...
  Advantage : still defined in SimDM

- Data
  From download or cutout
  Proposition : several formats as VO-Table, HDF5, FITS, ASCII, ...

Example : observations towards HD 102065

A&A 483, 485–493 (2008)
DOI: 10.1051/0004-6361:20078374

Astronomy & Astrophysics

Modeling of diffuse molecular gas applied to HD 102065 observations

C. Nehmé[1], J. Le Bourlot[1], F. Boulanger[2], G. Pineau des Forêts[2,5], and C. Gry[3,4]

[1] LUTH, UMR 8102, CNRS, Université Paris 7 and Observatoire de Paris, Place J. Janssen, 92195 Meudon, France
e-mail: Jacques.Lebourlot@obspm.fr
[2] Institut d'Astrophysique Spatiale, UMR 8617, CNRS, Université Paris-Sud, Bât. 121, 91405 Orsay Cedex, France
[3] Laboratoire d'Astrophysique de Marseille, UMR 6110, CNRS, Université de Provence, 38 rue Frédéric Joliot-Curie, 13388 Marseille Cedex 13, France
[4] European Space Astronomy Center, RSSD, PO Box 50727, 28080 Madrid, Spain
[5] LERMA, UMR 8112, CNRS, Observatoire de Paris, France

Received 28 July 2007 / Accepted 23 January 2008

ABSTRACT

Aims. We model a diffuse molecular cloud present along the line of sight to the star HD 102065. We compare our modeling with observations to test our understanding of physical conditions and chemistry in diffuse molecular clouds.
Methods. We analyze an extensive set of spectroscopic observations which characterize the diffuse molecular cloud observed toward HD 102065. Absorption observations provide the extinction curve, H_2, C I, CO, CH, and CH^+ column densities and excitation. These data are complemented by observations of C^+, CO and dust emission. Physical conditions are determined using the Meudon PDR model of UV illuminated gas.
Results. We find that all observational results, except column densities of CH, CH^+ and H_2 in its excited ($J \geq 2$) levels, are consistent with a cloud model implying a Galactic radiation field ($G \sim 0.4$ in Draine's unit), a density of 80 cm^{-3} and a temperature (60–80 K) set by the equilibrium between heating and cooling processes. To account for excited ($J \geq 2$) H_2 levels column densities, an additional component of warm (~250 K) and dense ($n_H \geq 10^4$ cm^{-3}) gas within 0.03 pc of the star would be required. This solution reproduces the observations only if the ortho-to-para H_2 ratio at formation is ~1. In view of the extreme physical conditions and the unsupported requirement on the ortho-to-para ratio, we conclude that H_2 excitation is most likely to be accounted for by the presence of warm molecular gas within the diffuse cloud heated by the local dissipation of turbulent kinetic energy. This warm H_2 is required to account for the CH^+ column density. It could also contribute to the CH abundance and explain the inhomogeneity of the CO abundance indicated by the comparison of absorption and emission spectra.

Key words. astrochemistry – ISM: clouds – ISM: molecules – ISM: structure – ISM: individual objects: Chamaeleon clouds – stars: individual: HD 102065

1. Introduction

Since the pioneering work of Black & Dalgarno (1977), observations of diffuse molecular clouds continue to motivate and challenge efforts to model the thermal balance and chemistry of interstellar gas illuminated by UV photons. Models allow observers to determine physical conditions from their data and observations contribute to models by quantifying physical processes of general relevance to studies of matter in space such as H_2 formation, photo-electric heating, and cosmic ray ionization.

Many models of well characterized lines of sight have been presented (e.g. in the last years: Zsargó & Federman 2003; Le Petit et al. 2004; Shaw et al. 2006). They are successful in reproducing many observables apart from some molecular abundances, most conspicuously CH^+, which points to out-of-equilibrium chemistry. This molecular ion, and several of the molecular species commonly observed in diffuse molecular clouds such as CH, OH and HCO^+ may be produced by MHD shocks (Draine & Katz 1986; Pineau des Forêts et al. 1986; Flower & Pineau des Forêts 1998), and small scale vortices (Joulain et al. 1998; Falgarone et al. 2006) where H_2 is heated by the localized dissipation of the gas turbulent kinetic energy. Turbulent transport between the cold and warm neutral medium may also significantly impact the chemistry of diffuse clouds (Lesaffre et al. 2007).

Independently of gas chemistry, the presence of H_2 at higher temperatures than that set by UV and cosmic-rays heating of diffuse molecular clouds, may be probed through observations of the H_2 level populations (Cecchi-Pestellini et al. 2006). A correlation between CH^+ and rotationally excited H_2 was found by Lambert & Danks (1986) using Copernicus observations. Falgarone et al. (2005) reported the detection of the S(0) to S(3) H_2 lines in a line of sight towards the inner Galaxy away from star forming regions. They interpret their observation as evidence for traces of warm molecular gas in the diffuse interstellar medium. But the interpretation of the wealth of H_2 observations provided by the FUSE satellite is still a matter of debate. Gry et al. (2002) modeled FUSE H_2 observations of three stars in Chamaeleon using the Meudon Photon Dominated Regions (PDR) model (Le Bourlot et al. 1993). They show that the model cannot account for H_2 column densities in rotational states with $J > 2$. A larger sample of H_2 FUSE observations (Tumlinson et al. 2002; Gillmon et al. 2005; Wakker 2006), including 2 of the 3 Chamaeleon lines of sight of Gry et al. (2002), have been analyzed on the basis of model calculations presented by Browning et al. (2003). Their model, like other PDR models, takes into

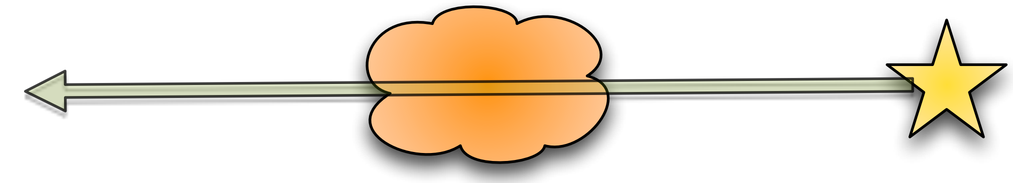


Table 1. Observational constraints and best model results. Upper part are constraints used in Fig. 2, lower part compares unconstrained observations and results. Number in parentheses are powers of 10.

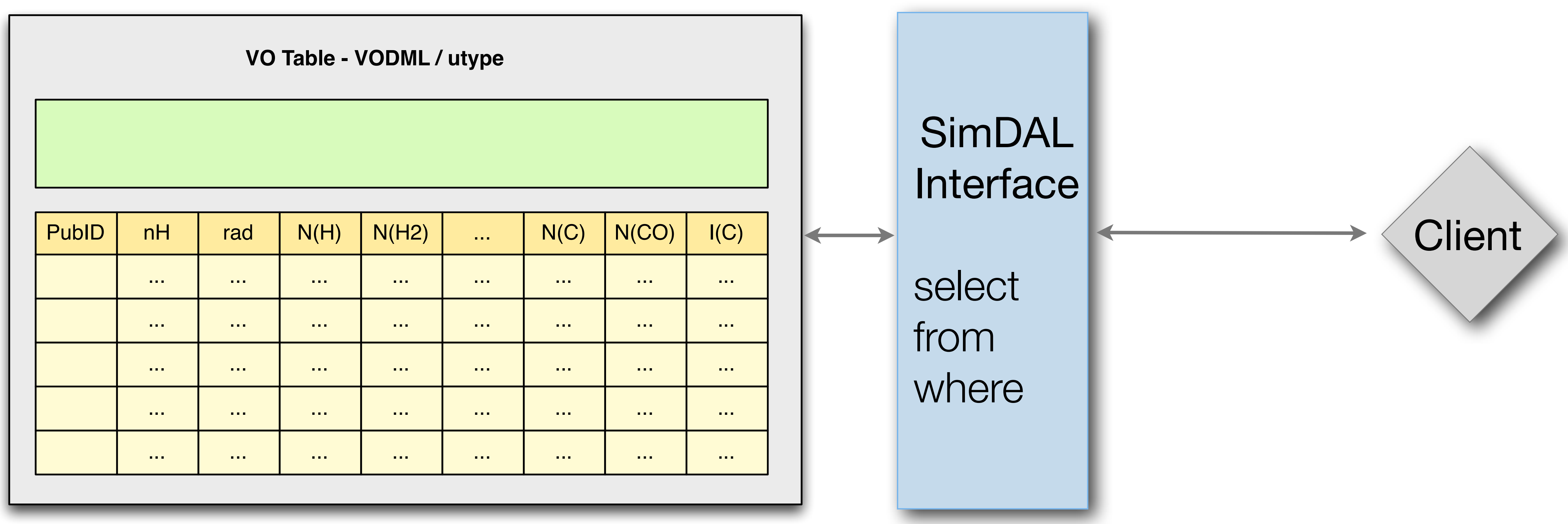
	X^{mod}	X^{obs}	σ_{obs}
$N(CO)/N(H_2)$	1.5 (−7)	1.6 (−7)	$\pm^{0.2(-7)}_{0.15(-7)}$
$N(C\,I)/N_H$	5.8 (−7)	6.0 (−7)	±1.5 (−7)
$N(C\,I^*_{J=1})/N(C\,I)$	0.17	0.16	±0.07
$N(C\,I^{**}_{J=2})/N(C\,I)$	0.03	0.024	±0.01
$f_{H_2} = \frac{2N(H_2)}{N(H)+2N(H_2)}$	0.9	0.69	±0.12
$N(H_2^o)/N(H_2^p)$	0.73	0.7	±0.12
$I(C^+)$ (erg/s cm^2 sr)	2.0 (−6)	2.8 (−6)	±0.85 (−6)
$N(CH)/N(H_2)$	8.4 (−9)	1.85 (−8)	±0.3 (−8)
$N(CN)/N(H_2)$	1.2 (−10)	<1.5 (−9)	
$N(C_2)/N(H_2)$	3.6 (−8)	<3.5 (−8)	
$N(CO_{J=0})/N(H_2)$	9.0 (−8)	9.6 (−8)	$\pm^{1.4(-8)}_{1.7(-8)}$
$N(CO_{J=1})/N(H_2)$	5.1 (−8)	6.2 (−8)	$\pm^{1.5(-8)}_{1.2(-8)}$
$N(CO_{J=2})/N(H_2)$	3.7 (−9)	<7.3 (−9)	

Observed quantities

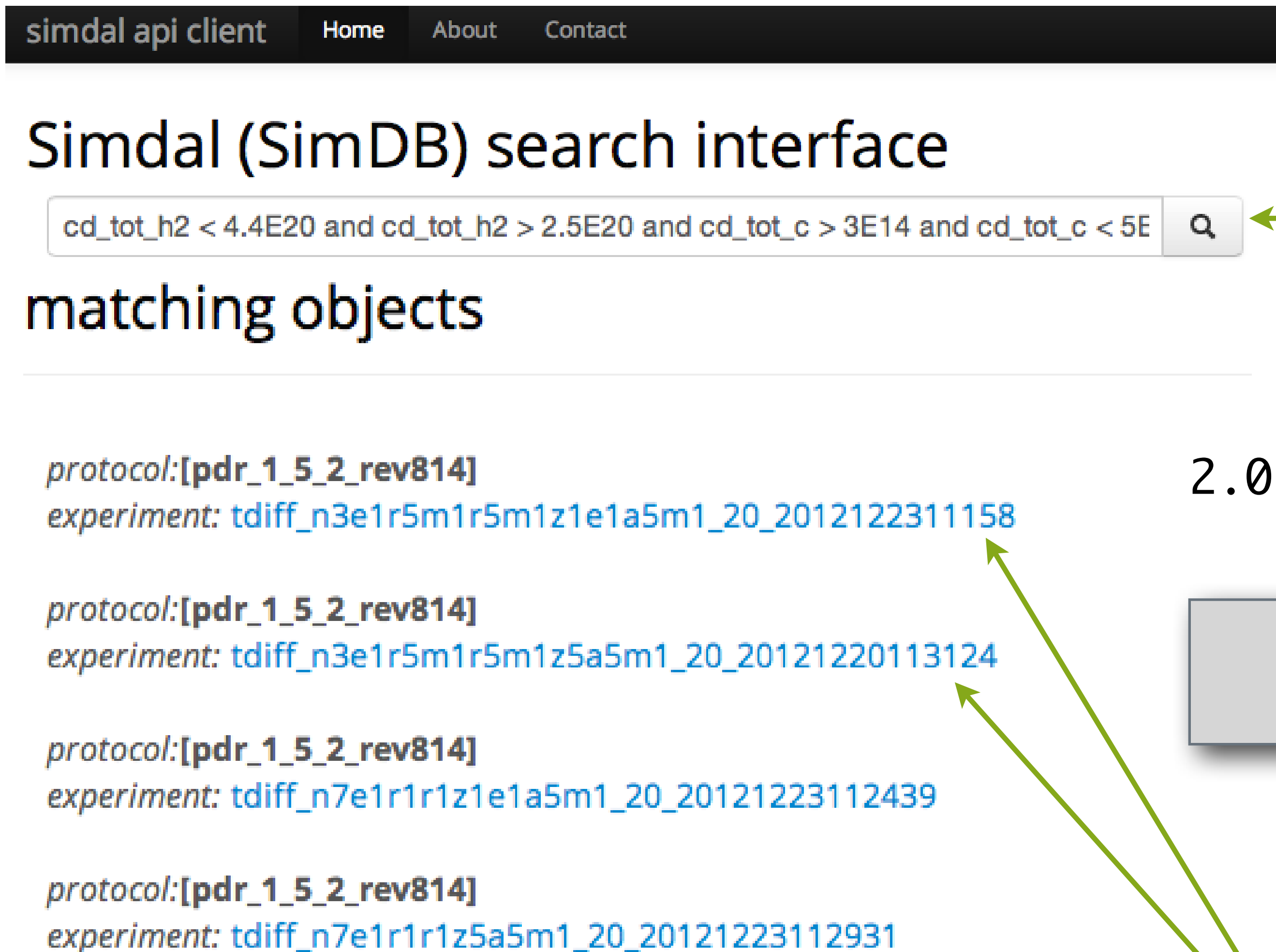
- Column densities H_2, C, CO
- H_2 populations in rotationnal levels
- Line intensity of C^+ at 158 μm

Example : observations towards HD 102065

Metadata are stored in virtual tables



Example : observations towards HD 102065



Query:

$2.4\ 10^{20} < N(H2) < 4.4\ 10^{20}\ cm^{-2}$

$3.0\ 10^{14} < N(C) < 5\ 10^{14}\ cm^{-2}$

$N(CO) < 1.0\ 10^{14}\ cm^{-2}$

$2.0\ 10^{-6} < I(C+,\ 158\ \mu m) < 4\ 10^{-6}\ erg\ s^{-1}\ cm^{-2}\ sr^{-1}$

Models in agreement with query

Datalink

Example : observations towards HD 102065

simdal api client   Home   About   Contact

Simdal (SimDB) search interface

cd_tot_h2 < 4.4E20 and cd_tot_h2 > 2.5E20 and cd_tot_c > 3E14 and cd_tot_c < 5E

matching objects

*protocol:***[pdr_1_5_2_rev814]**
experiment: tdiff_n3e1r5m1r5m1z1e1a5m1_20_2012122311158

*protocol:***[pdr_1_5_2_rev814]**
experiment: tdiff_n3e1r5m1r5m1z5a5m1_20_20121220113124

*protocol:***[pdr_1_5_2_rev814]**
experiment: tdiff_n7e1r1r1z1e1a5m1_20_20121223112439

*protocol:***[pdr_1_5_2_rev814]**
experiment: tdiff_n7e1r1r1z5a5m1_20_20121223112931

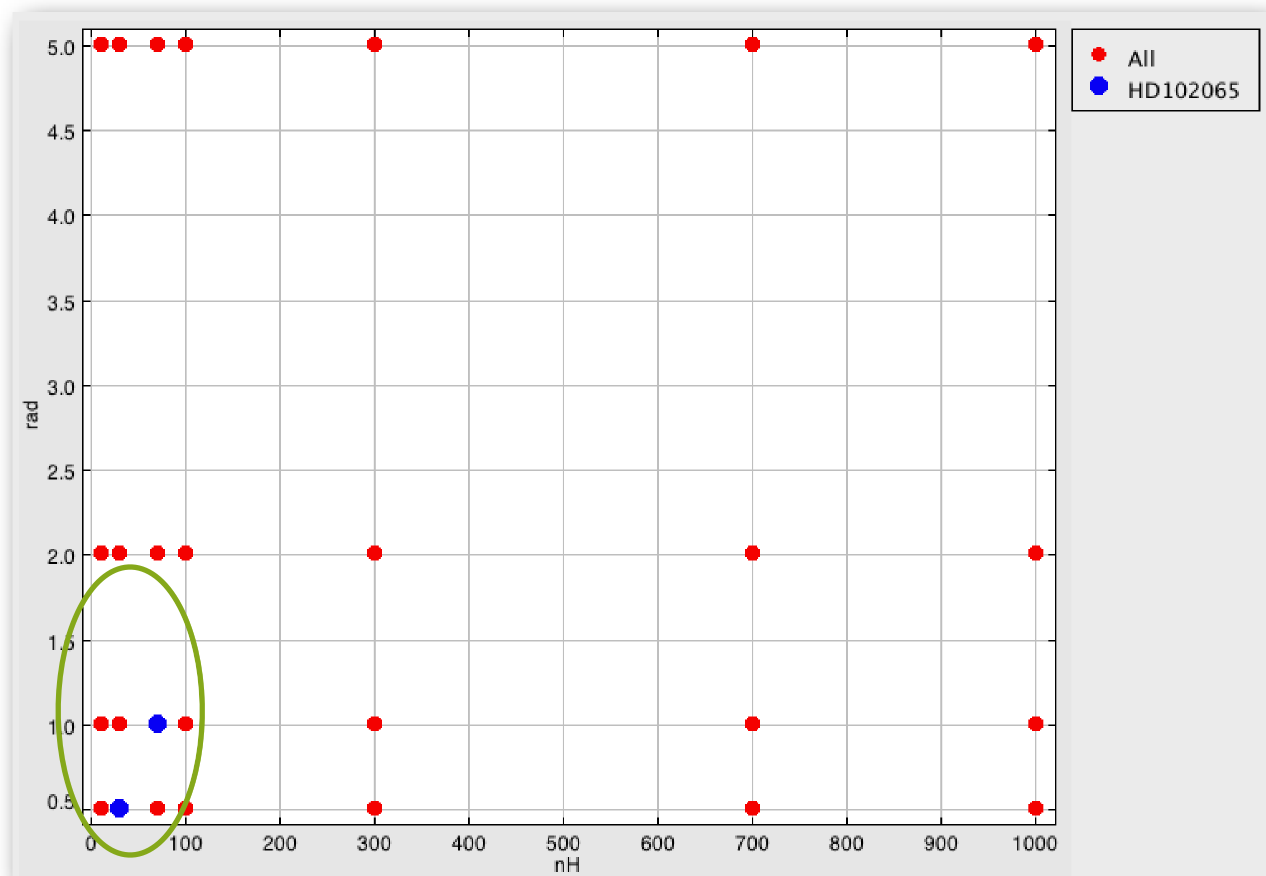
Query:

$2.4\ 10^{20} < N(H2) < 4.4\ 10^{20}\ cm^{-2}$

$3.0\ 10^{14} < N(C) < 5\ 10^{14}\ cm^{-2}$

$N(CO) < 1.0\ 10^{14}\ cm^{-2}$

$2.0\ 10^{-6} < I(C+,\ 158\ \mu m) < 4\ 10^{-6}\ erg\ s^{-1}\ cm^{-2}\ sr^{-1}$



Interprétation des observations vers HD 102065

simdal api client | Home | About | Contact

Simdal (SimDB) search interface

cd_tot_h2 < 4.4E20 and cd_tot_h2 > 2.5E20 and cd_tot_c > 3E14 and cd_tot_c < 5E

matching objects

protocol:[pdr_1_5_2_rev814]
experiment: tdiff_n3e1r5m1r5m1z1e1a5m1_20_20121223111158

protocol:[pdr_1_5_2_rev814]
experiment: tdiff_n3e1r5m1r5m1z5a5m1_20_20121220113124

protocol:[pdr_1_5_2_rev814]
experiment: tdiff_n7e1r1r1z1e1a5m1_20_20121223112439

protocol:[pdr_1_5_2_rev814]
experiment: tdiff_n7e1r1r1z5a5m1_20_20121223112931

Cutout

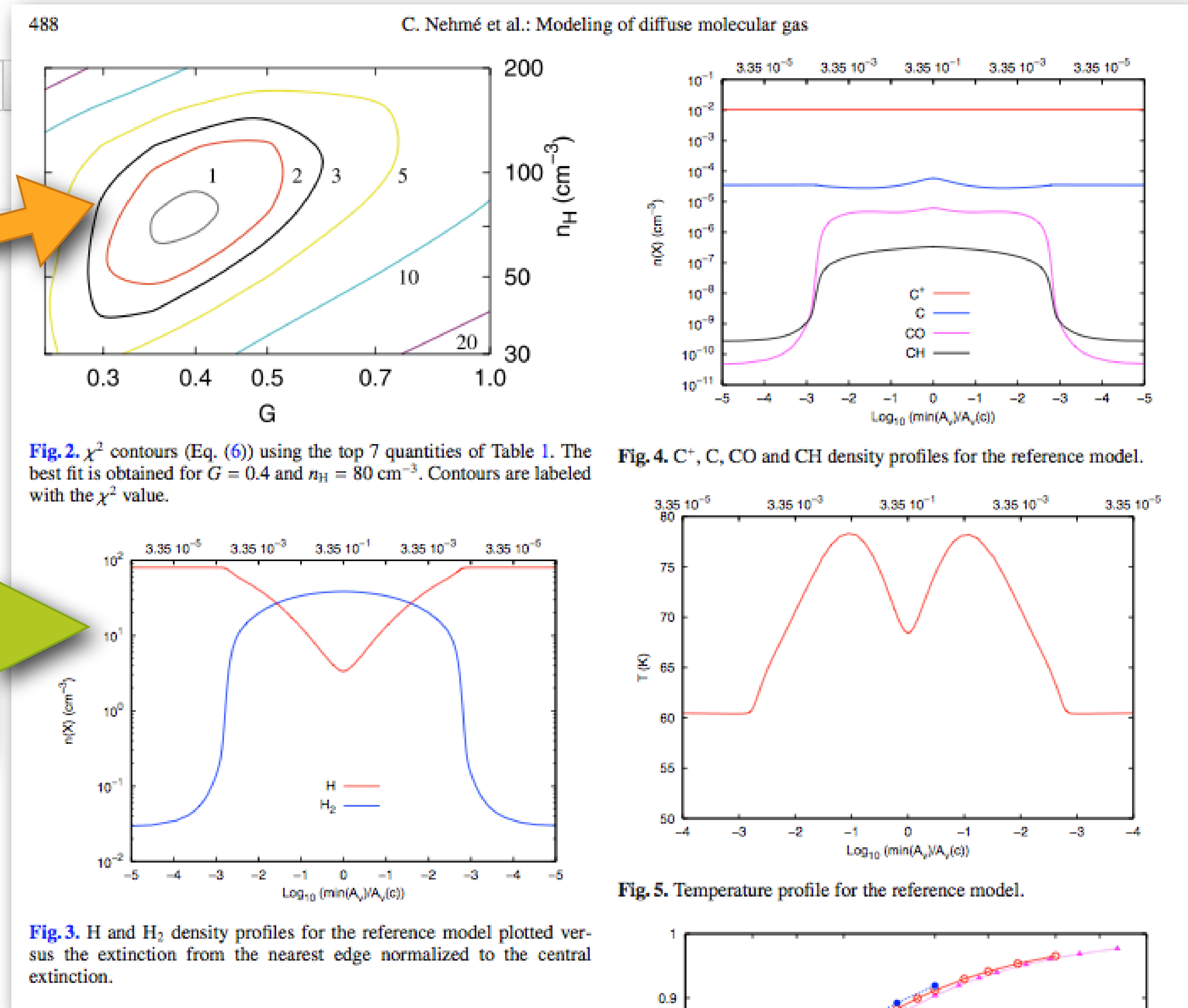

488 C. Nehmé et al.: Modeling of diffuse molecular gas

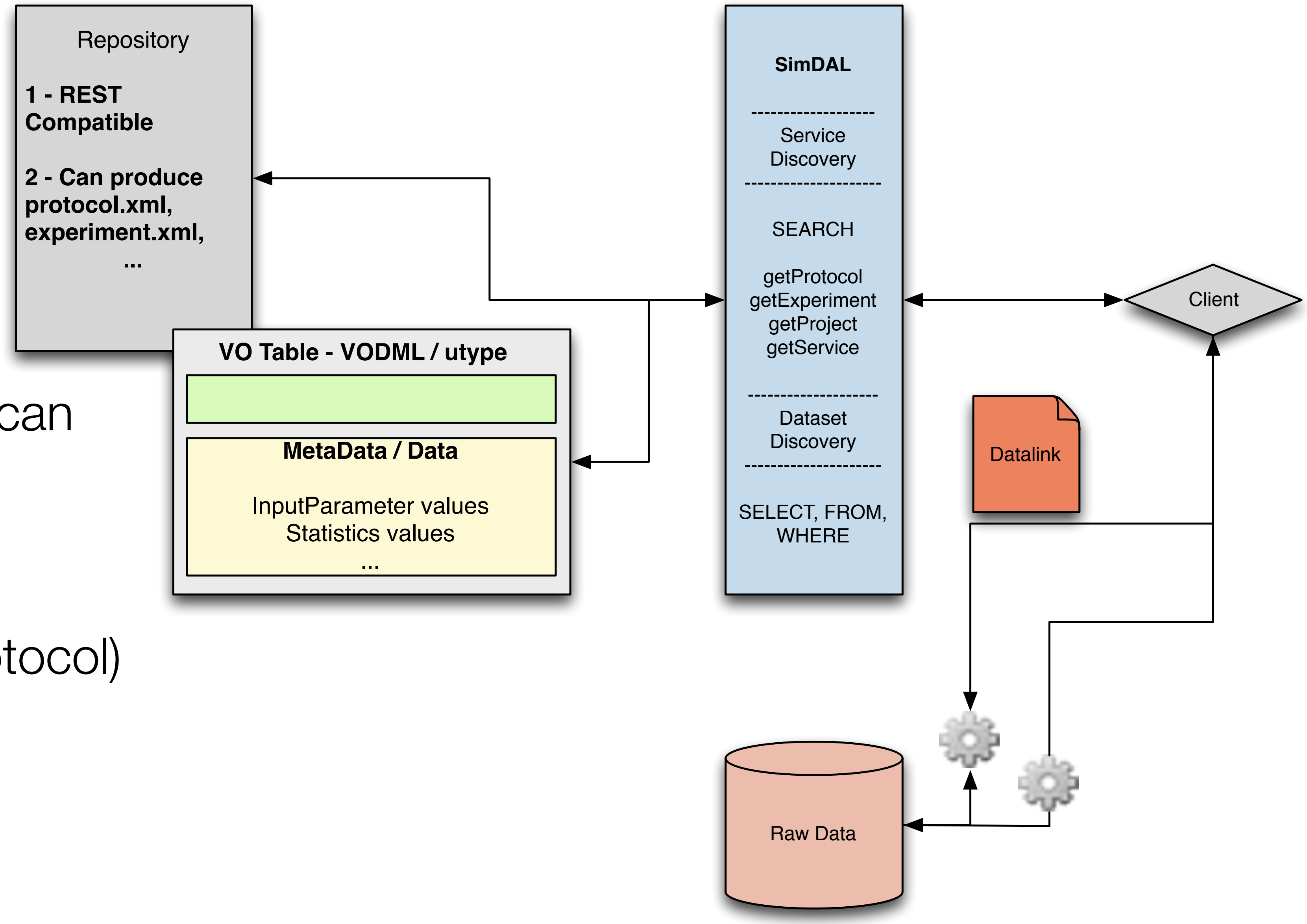
Fig. 2. χ^2 contours (Eq. (6)) using the top 7 quantities of Table 1. The best fit is obtained for $G = 0.4$ and $n_H = 80$ cm^{-3}. Contours are labeled with the χ^2 value.

Fig. 3. H and H_2 density profiles for the reference model plotted versus the extinction from the nearest edge normalized to the central extinction.

Fig. 4. C^+, C, CO and CH density profiles for the reference model.

Fig. 5. Temperature profile for the reference model.

Discussions



Queries

Repository : REST compliant and can produce .xml (SimDM)

- definition of a few functions (example : get description of protocol)
- Select, From, Where

Exchanged files

- protocol.xml, experiment.xml, project.xml, service.xml
- Data in any format : VO-Table, HDF5, FITS, ASCII

Cutout

- Is a generic cutout planned (any-dimensional axis) ?
- how to specify the list of quantities to retrieve (and of possibilities) ?