



Semantics & Theory

Franck Le Petit, Nicolas Moreau
Zakaria Méliani , David Languignon, Emeric Bron



SimDM & Semantics

Simulation Data Model (SimDM) is a data model to describe numerical simulations

SimDM goal: help scientists to discover simulations through queries on various quantities

- simulated processes
- input parameters
- computed quantities (statistics on simulation results)

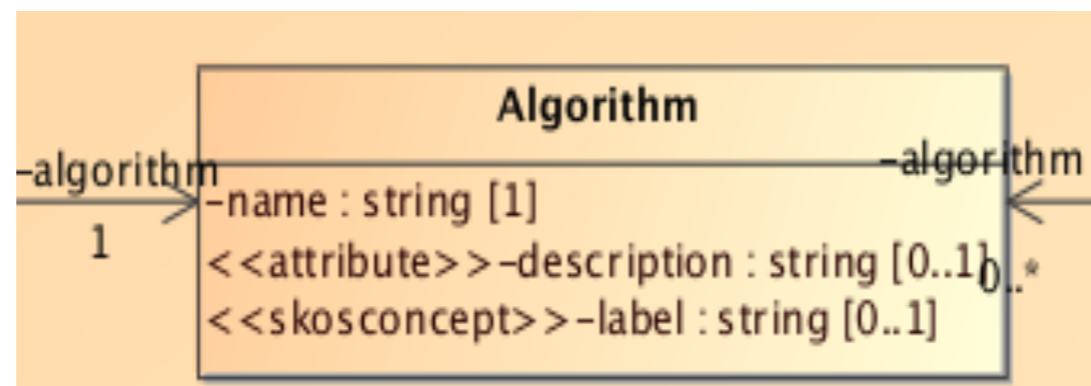
Numerical simulations are diverse

- SimDM is a meta-model: parameters, properties are not explicitly defined
- use of vocabularies to characterize quantities

SimDM + vocabularies

SimDM uses **SKOS** (Simple Knowledge Organization System) vocabularies

- some classes have an skosconcept attribute



Theory vocabularies

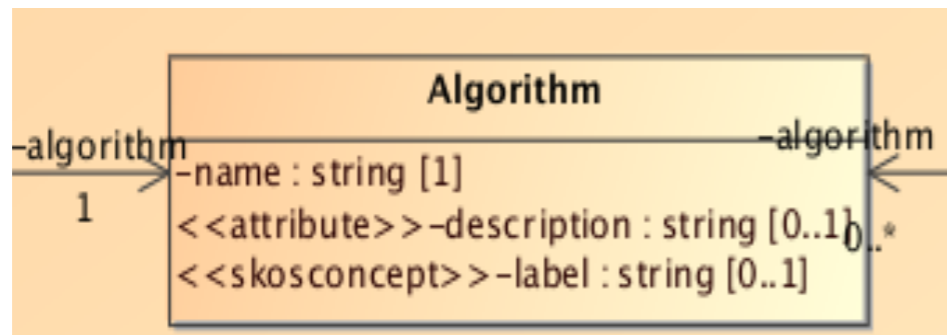
Theory SKOS vocabularies:

- concept URI
- PREF label
- ALT labels
- Narrower / Broader

<http://ivoa.net/rdf/theory/Algorithms#ForwardTimeCentralSpace>

Common name

Synonymes



<http://ivoa.net/rdf/theory/Algorithms#ForwardTimeCentralSpace>

PREF: Forward Time Central Space
ALT: FTCS
Broader: Finite difference
Related: Lax-Friedrichs

Forward-Time Central-Space

Finite difference method used to solve parabolic partial differential equations. The method is first-order, explicit and conditionally stable ("Computational Fluid Mechanics and Heat Transfer 2nd ed.", John C. Tannehill, Dale A. Anderson, Richard H. Pletcher, 1997).

<http://ivoa.net/rdf/theory/Algorithms#ForwardTimeCentralSpace>

AltLabels

FTCS (en)

Broader concepts

[Finite Difference](#)

Broader Transitive concepts

[Algorithm](#)

[Finite Difference](#)

Related concepts

[Lax-Friedrichs](#)

Common & specific vocabularies

Numerical simulations are diverse

- some concepts are specific to a code

Example: Choice between home-made algorithms to simulate a process

- some concepts do not have a precise definition

Example: Mean UV radiation field intensity (G_0): definition depends on the author

- complex to define proper nomenclature for some concepts

Example: Line intensity of the $H_2^{18}O$ between levels 1_{10} and 0_{00}

→ need for **2 kinds of vocabularies**

- **Official** vocabularies: contain the most common concepts (official concepts)
- **Specific** vocabularies: contain uncommon concepts (specific to a team / service)
 - do not favour interoperability but we need them
 - when a specific concept becomes broadly used it can go in the official vocabularies

Theory vocabularies

Official vocabularies

SimDM requires vocabularies for:

- Algorithms
- Astronomical objects
- Data object types
- Physical properties
- Physical quantities

Controlled vocabularies

- versionning
- Mailing list to suggest new concepts:
support.votheory@obspm.fr.

Specific vocabularies

Projects can need specific vocabularies.

- concepts specific to a code / to some simulations
- complex concepts that would need standardisation to know how to manage them in official vocabularies

Example: ISMService - <http://ism.obspm.fr>

Need to describe each atomic and molecular line + many other concepts

→ specific vocabulary with ~ **300 000 concepts**

Official vocabularies

- **Algorithms** ~ 122 concepts
Runge-Kutta, Burlish-Stoer, ...
 - **Physical processes** ~ 128 concepts
Ex: turbulence, gravitation, ...
 - **Physical quantities** ~ 131 concepts
Ex: Velocity, Mass, ...
 - **Data Objects Types** ~ 17 concepts
Ex: mesh cell, ...
 - **Astronomical Objects** ~ 236 concepts
→ *Dégradation de l'ontologie (OWL) en SKOS*
- Total: ~**634 concepts**

Discovering concepts for publication of simulations

Need convenient GUI so scientists can find URIs

- discover concepts
- navigate between concepts

<http://votheory.obspm.fr>

Official vocabularies

Specific vocabularies

The screenshot shows the IVOA theory website interface. At the top, there are navigation links for 'Home', 'Search concepts', and 'Help'. Below this, there are two tabs: 'IVOA vocabularies' and 'Specific vocabularies'. A paragraph explains that these are high-level metadata for astrophysical data. A dropdown menu is set to 'Algorithms', with a description: 'Vocabulary that defines numerical methods in use to obtain the data results.' Below this is a 'Concepts' section with a 'Quick search' input field. A large grid of concept tags is displayed, including '3+1 Formalism', '8-Wave Scheme', 'Accelerated Lambda Iteration', 'Adaptive Mesh Refinement', 'Advection Upstream Splitting Method', 'Algorithm', 'Alternating Direction Implicit', 'BiConjugate Gradient', 'BiConjugate Gradient Stabilized', 'Block Based AMR', 'Bulirsch-Stoer', 'Cell Based AMR', 'Cell Centred', 'Central Difference Scheme', 'Chebyshev Iteration', 'Conjugate Gradient Method', 'Conjugate Gradient Squared Method', 'Constrained Transport', 'Coupled Escaped Probability', 'Crank-Nicolson', 'Discontinuous Galerkin', 'Discontinuous Galerkin methods', 'Escape Probability', 'Euler', 'Exact Radiative Transfer Method', 'Exact Riemann Solver', 'Extended Finite Element Method', 'Fast-Multipole Method', 'Finite Difference', 'Finite Element', and 'Finite Volume'. On the right, a detailed view for 'Coupled Escaped Probability' is shown, including a citation: 'Exact method for line radiative transfer ("A new exact method for line radiative transfer", Elitzur, M., & Asensio Ramos, A. 2006, MNRAS, 365, 779).'. Below the citation is a URL: <http://ivoa.net/rdf/theory/Algorithms#CoupledEscapedProbability>. This URL is linked to a yellow oval labeled 'URI'. The detailed view also includes sections for 'AltLabels' (containing 'CEP (en)'), 'Broader concepts' (containing 'Algorithm' and 'Escape Probability'), and 'Broader Transitive concepts' (containing 'Algorithm'). An arrow from the 'Algorithm' link in the 'Broader concepts' section points to a yellow oval labeled 'Relations'.

Easy discovery of concepts

URI

Relations

Discovering concepts for publication of simulations

Reminder of Theory I.G. requirements concerning vocabularies

- ① **Persistent URIs** for concepts of official vocabularies
 - we asked at Shanghai InterOp to have ivoa.net URIs instead of purl.obspm.fr
- ② Store official controlled Theory **vocabularies on IVOA webpages**
 - officialisation of the vocabularies
 - we just need to publish XML files at the level of IVOA
- ③ Any simulation publisher can use its own **specific vocabularies**
 - so there are other Theory vocabularies than the ones at IVOA
with recommendation to use official concepts whenever possible for interoperability reasons
- ④ **Convenient GUI** to discover concepts - example: <http://votheory.obspm.fr>

Difficulties with the Semantics W.G.

Theory I.G. tries to fit in Semantics requirements but some difficulties

① Unilateral **changes between InterOps conclusions and implementation**

Exemple: Conclusion of Victoria InterOp:

<https://wiki.ivoa.net/internal/IVOA/InterOpMay2018PlenaryTCG/SemanticsClosingInteropMay2018.pdf>

• **Vocabularies**

- Set up ivoa.net/vocabularies/ on VO main page
- Store the vocabularies

2018/07/30

and the implementation is: [ivoa.net/rdf/...](http://ivoa.net/rdf/)

② IVOA (and semantics group) define standards not the way to present standards

How does a scientist is expected to discover concepts in that ?

IVOA Vocabulary: VOTheory Algorithms

This is the description of the namespace
<http://www.ivoa.net/rdf/theory/Algorithms> as of 2019-02-27.

This vocabulary is not yet approved by the IVOA. This means that terms can still disappear without prior notice.

Concepts in this vocabulary are intended to be used to describe algorithms in SimDM classes. To suggest new concepts or submit corrections, contact support.vothery@obspm.fr.

Predicate	Label	Description	Broader	Narrower
#3plus1Formalism	3+1 Formalism	Method used to solve relativity equations. 3+1 Formalism is an approach to general relativity and to Einstein equations that relies on the slicing of the four-dimensional spacetime by three-dimensional surfaces (hypersurfaces)	#Algorithm	
#8WaveScheme	8-Wave Scheme Powel		#Algorithm	
#AcceleratedLambdaIteration	Accelerated Lambda Iteration ALI	Method used to solve radiative transfer problems.	#Algorithm	
#AdaptiveMeshRefinement	Adaptive Mesh Refinement AMR	Adaptive mesh refinement (AMR) is a method of adapting the accuracy of a solution within certain precision.	#Algorithm	
#AdvectionUpstreamSplittingMethod	Advection Upstream Splitting Method	Method used to solve a general system of conservation equations.	#FiniteVolume	

Status of Theory vocabularies

- Vocabularies are **ready**
- URIs are in: <http://ivoa.net/rdf/theory/Algorithms#DiscontinuousGalerkin>
- Theory Group would like **to be publish its vocabularies on the IVOA Semantics webpage**

But issues with the semantics W.G.

- Need more concertation
- Browsing vocabularies with other GUI than the one defined by Semantics is refused by semantics

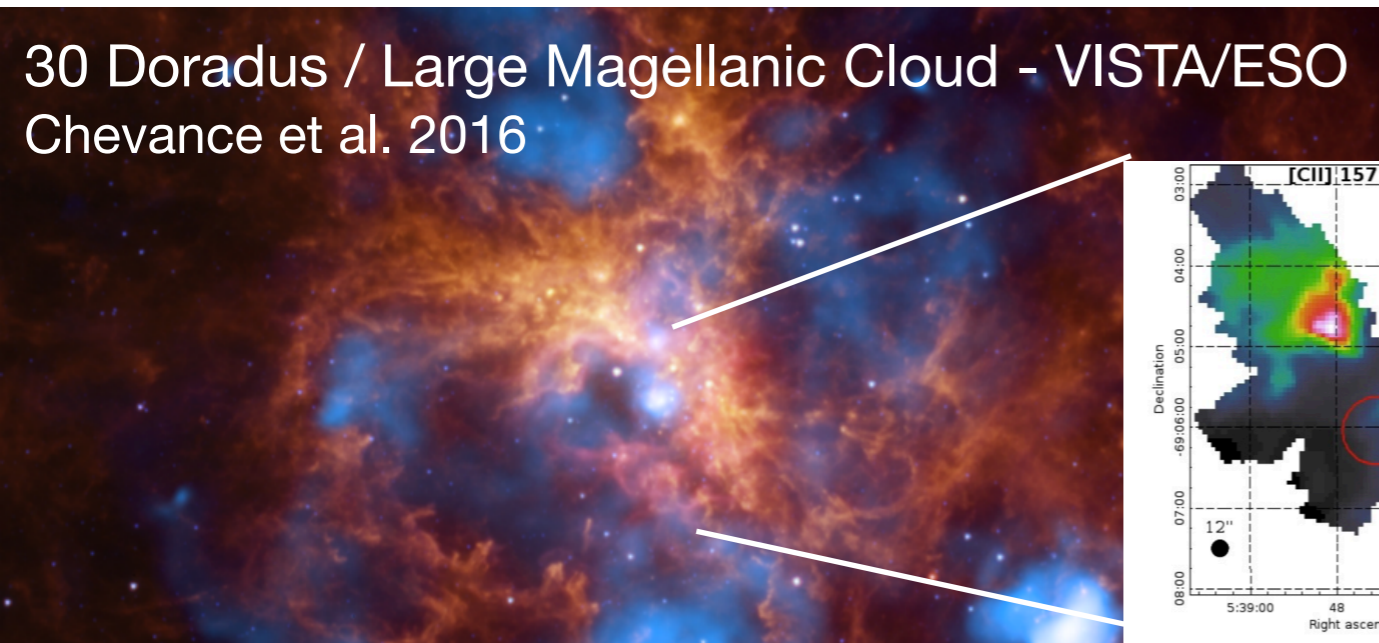
Solutions:

- ① we find an agreement with Semantics chair
- ② we publish our vocabularies on the Theory IVOA page
- ③ we publish our vocabulaires outside of IVOA

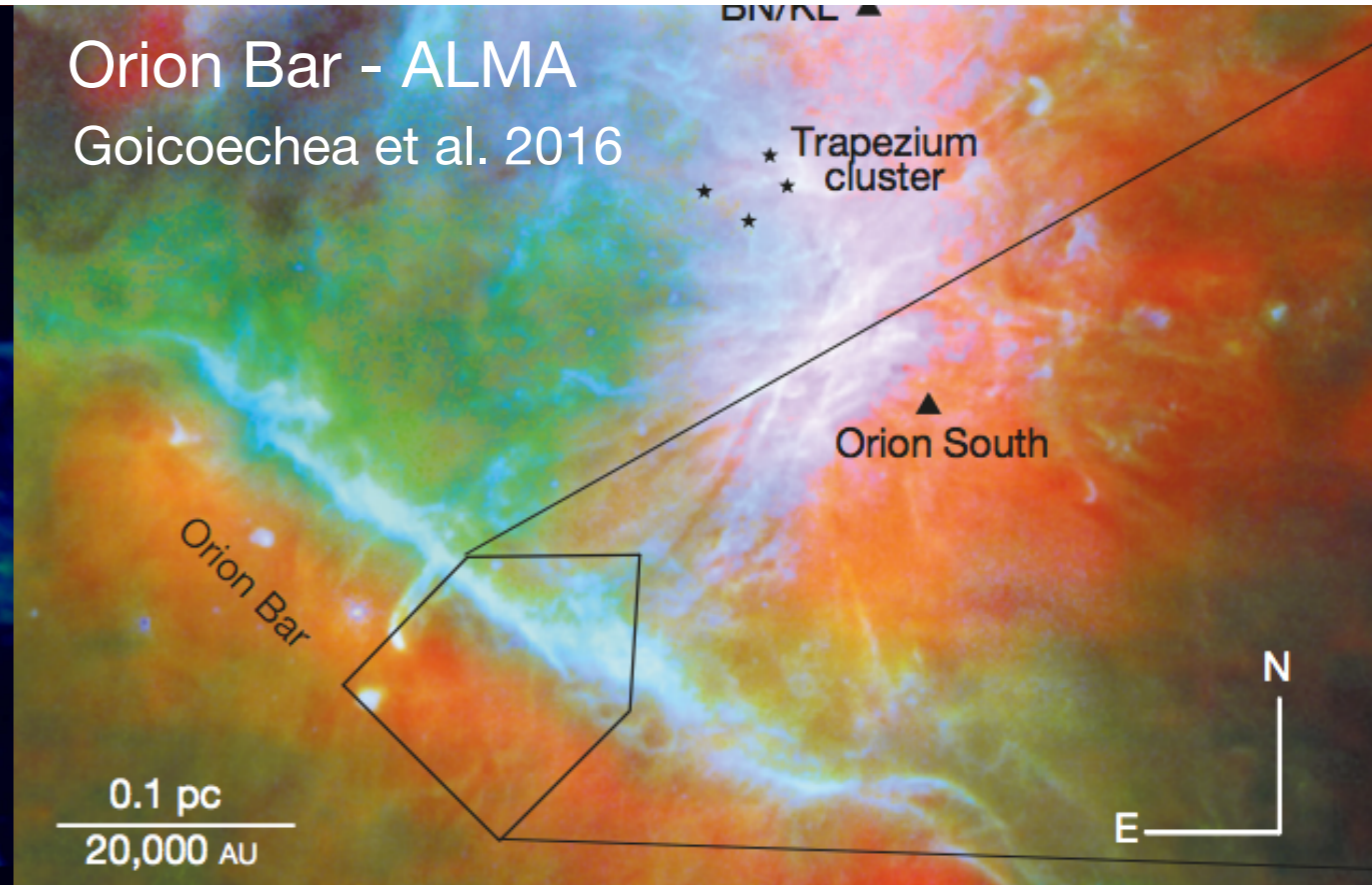
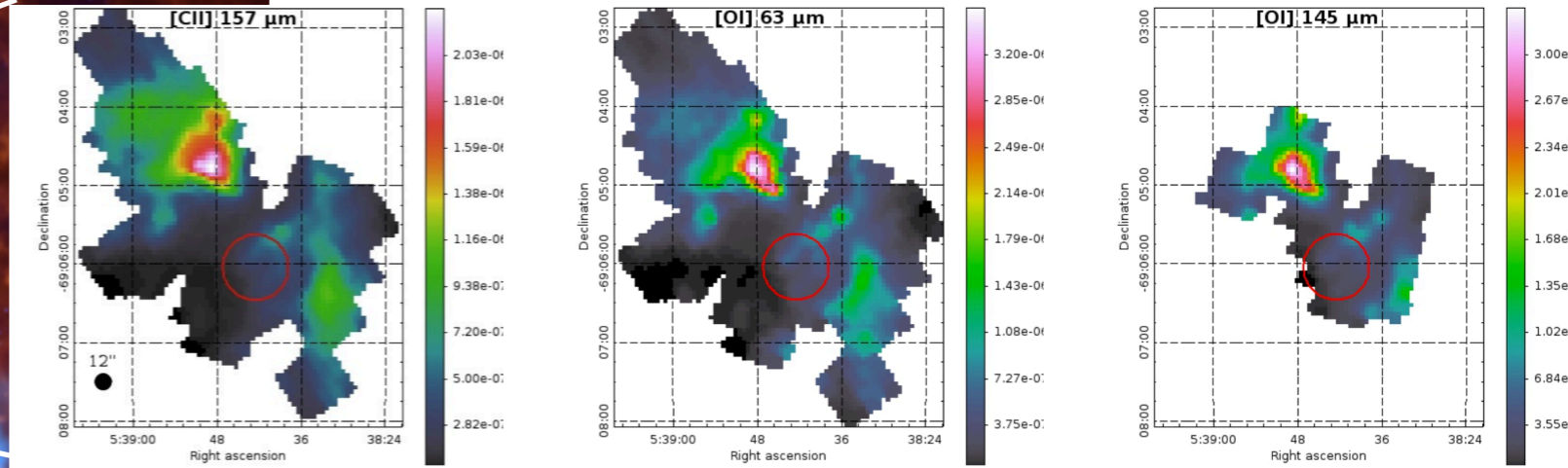
Example of vocabulary usage

or why do we need for Theory large vocabularies with relations and synonyms ?

Goal of ISMDB: interpret observations of interstellar gas emission



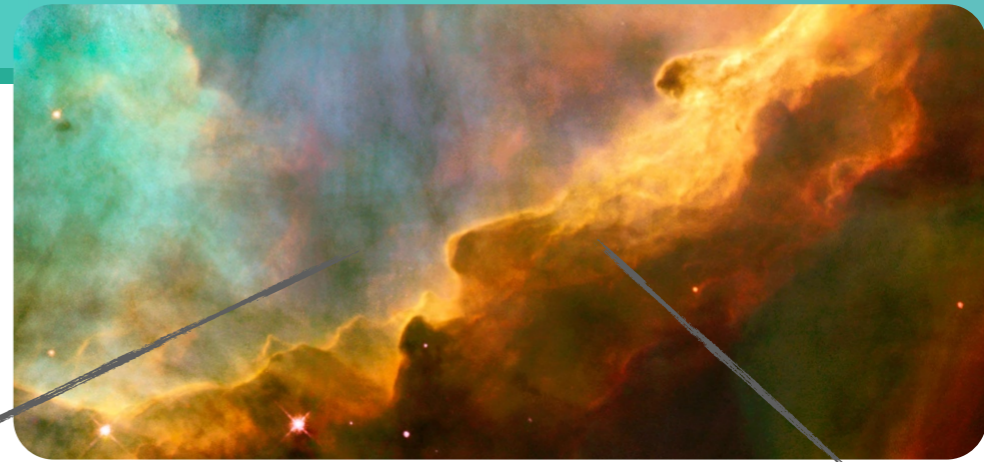
Herschel / PACS observations
C II 157 μm , O I 63 μm , O I 145 μm



Application of vocabularies: ISMDB

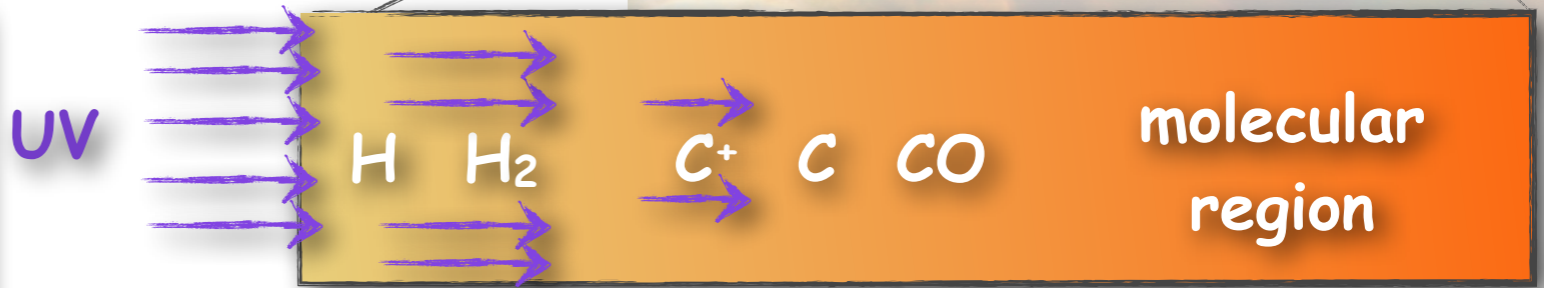
PDR codes:

- compute the atomic and molecular structure of interstellar clouds.
- analysis of physical and chemical processes

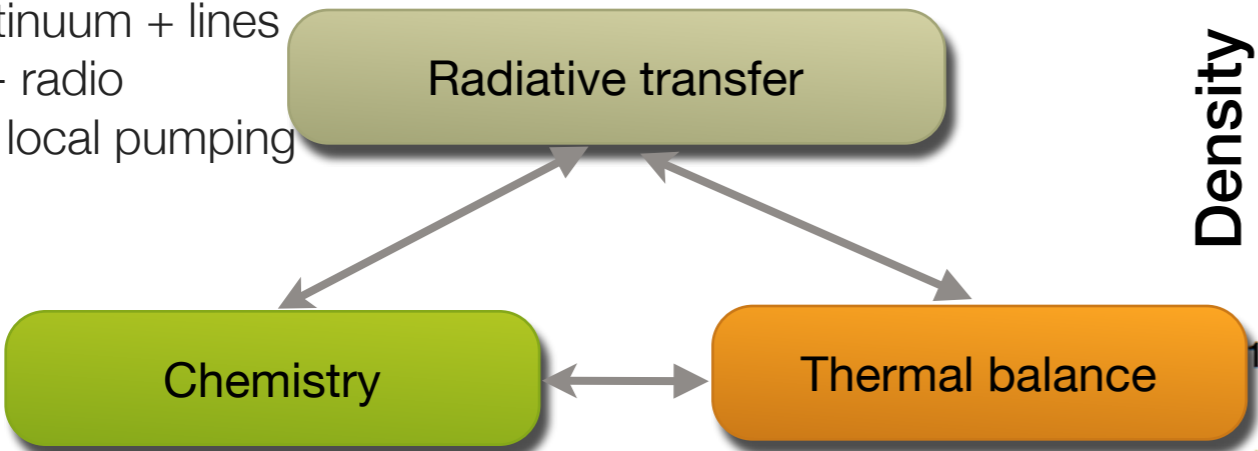


Outputs

- densities
- excitation
- gas & grains temperatures
- Intensities (H_2 , CO , H_2O , ...)
- Column densities

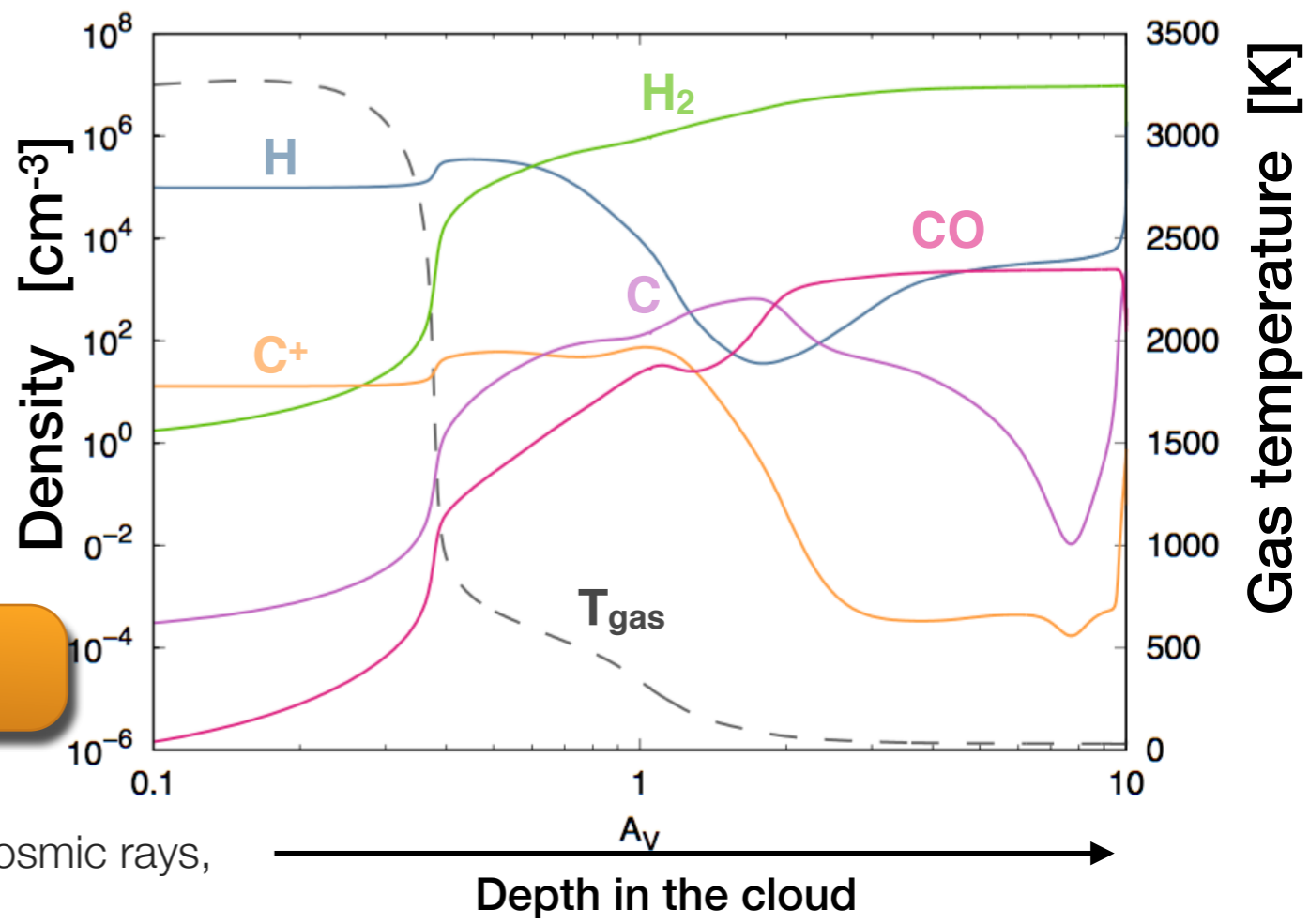


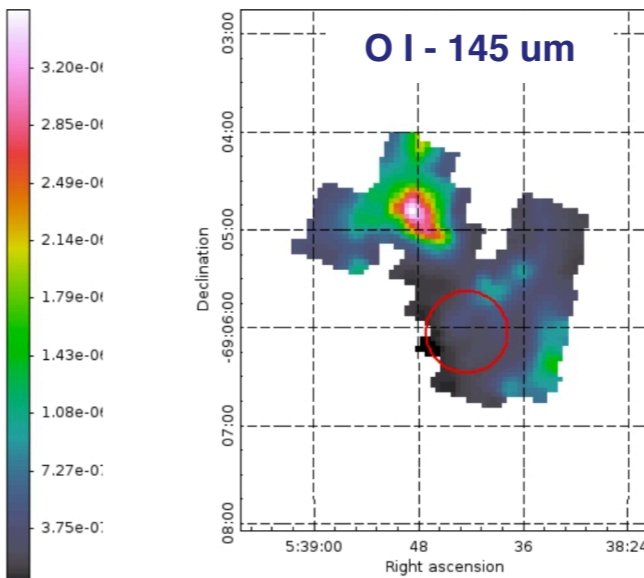
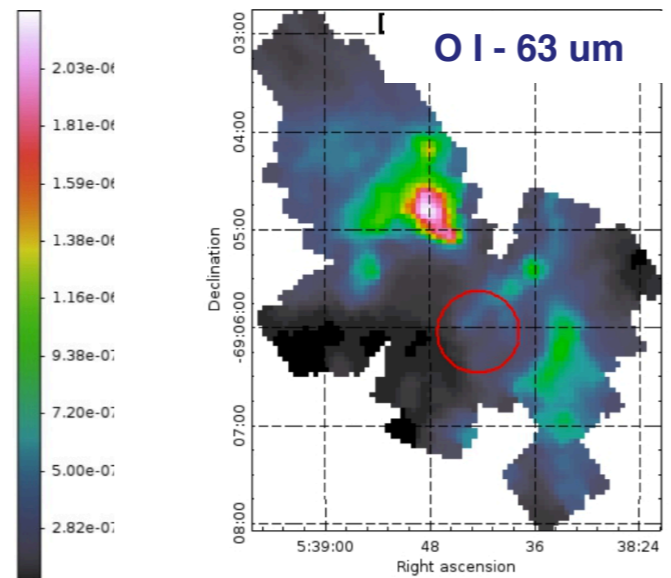
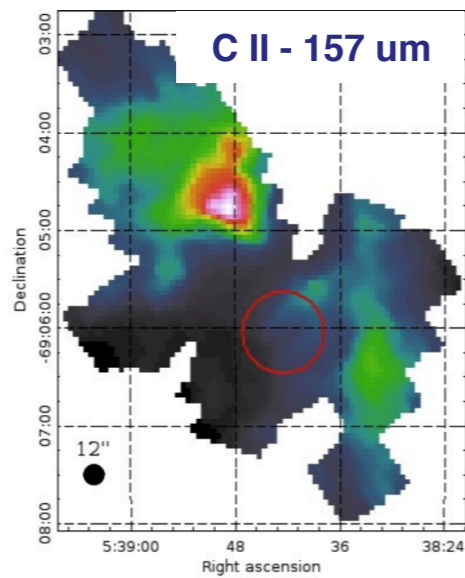
- continuum + lines
- UV - radio
- non local pumping



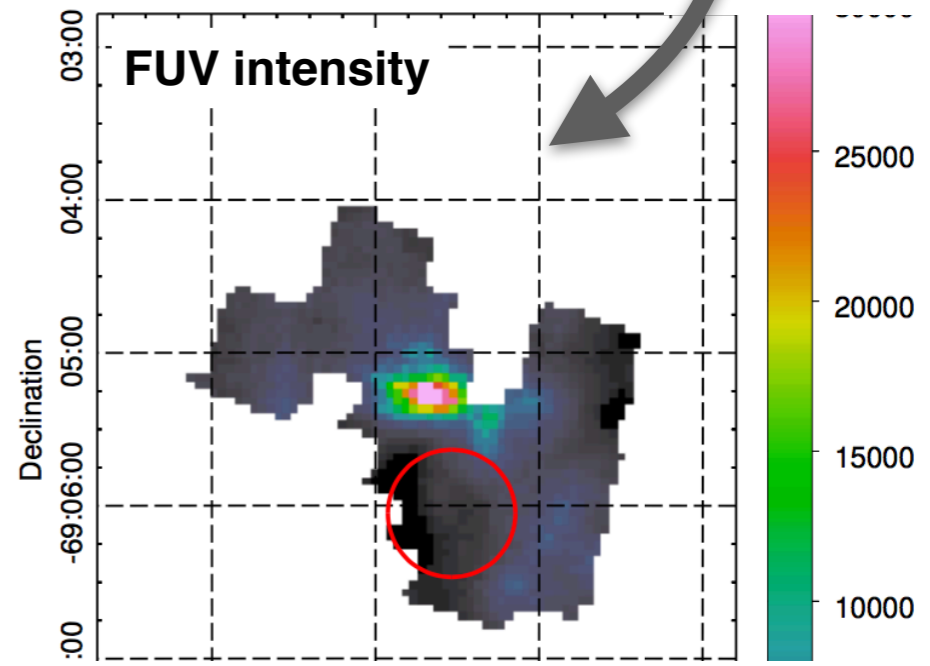
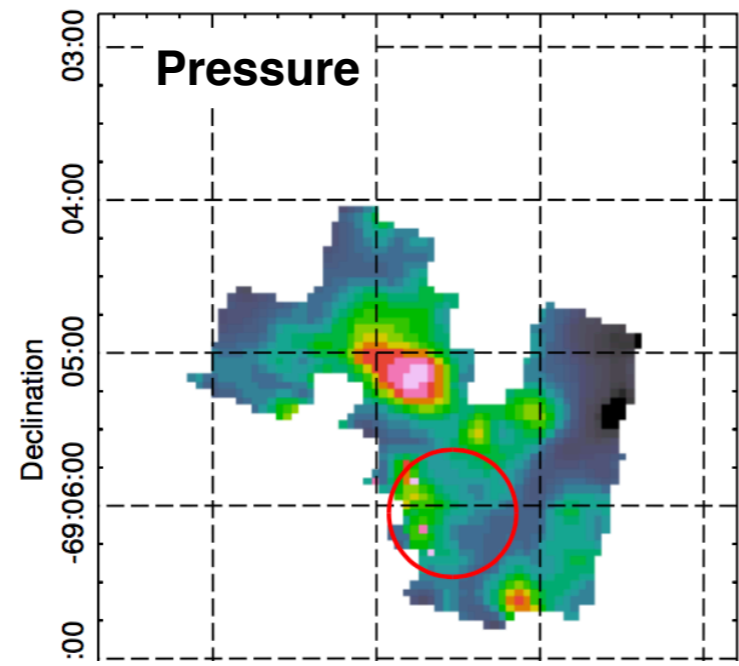
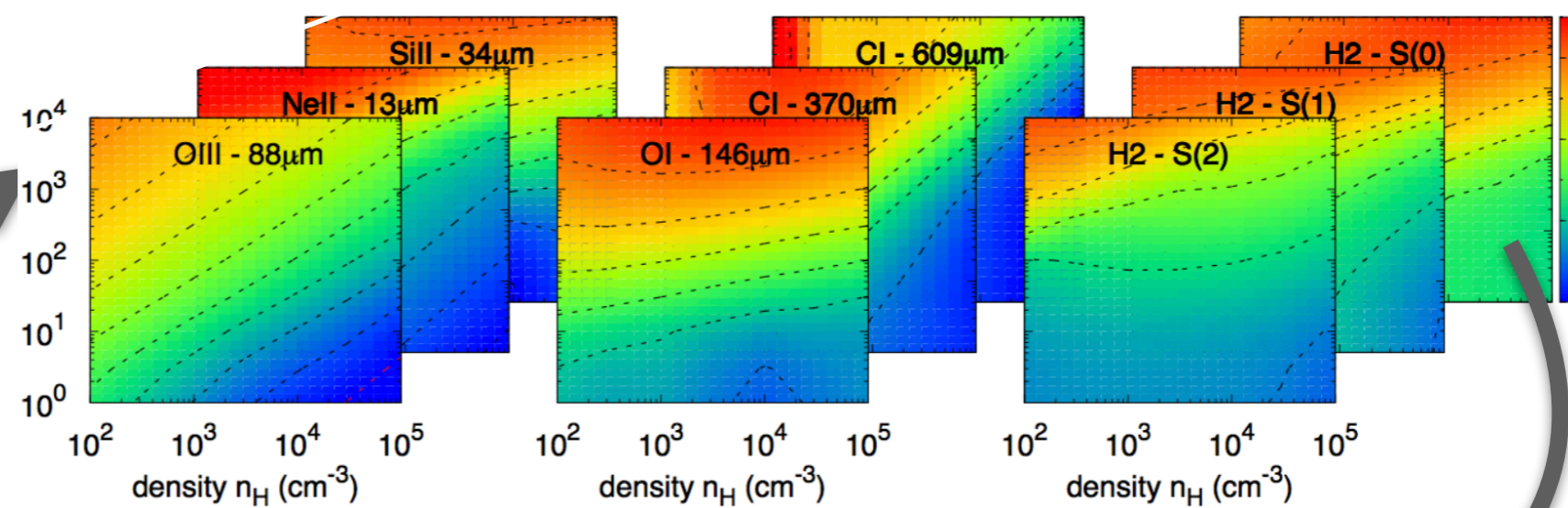
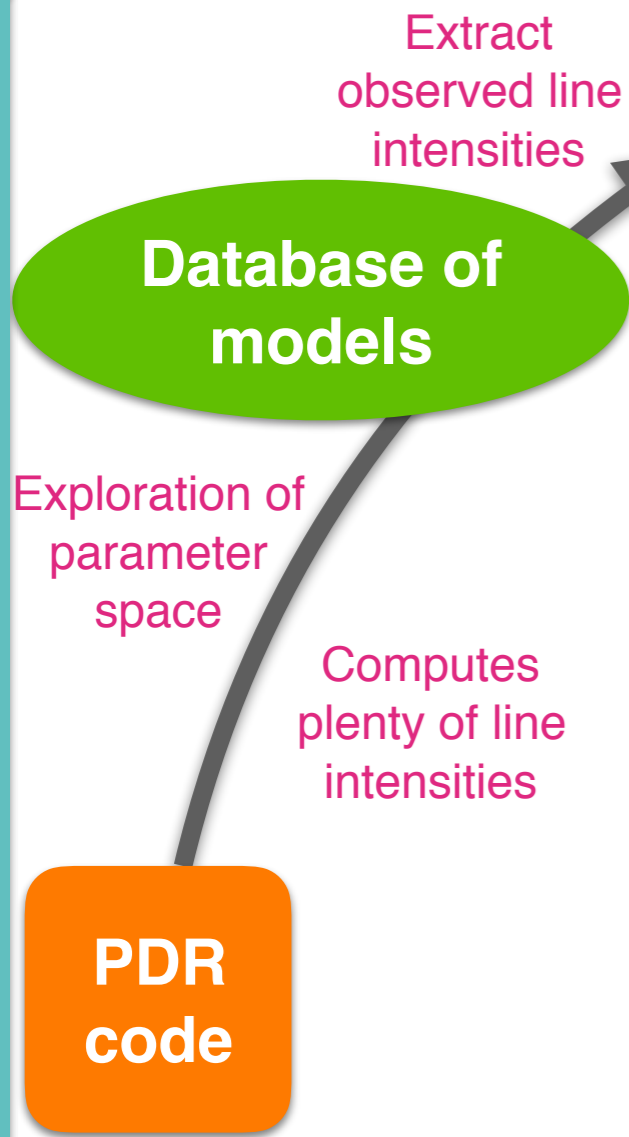
- several hundred species
- thousands reactions
- gas & grains

- Statistical equilibrium
- Photo-electric effect, cosmic rays, chemistry, ..., ...



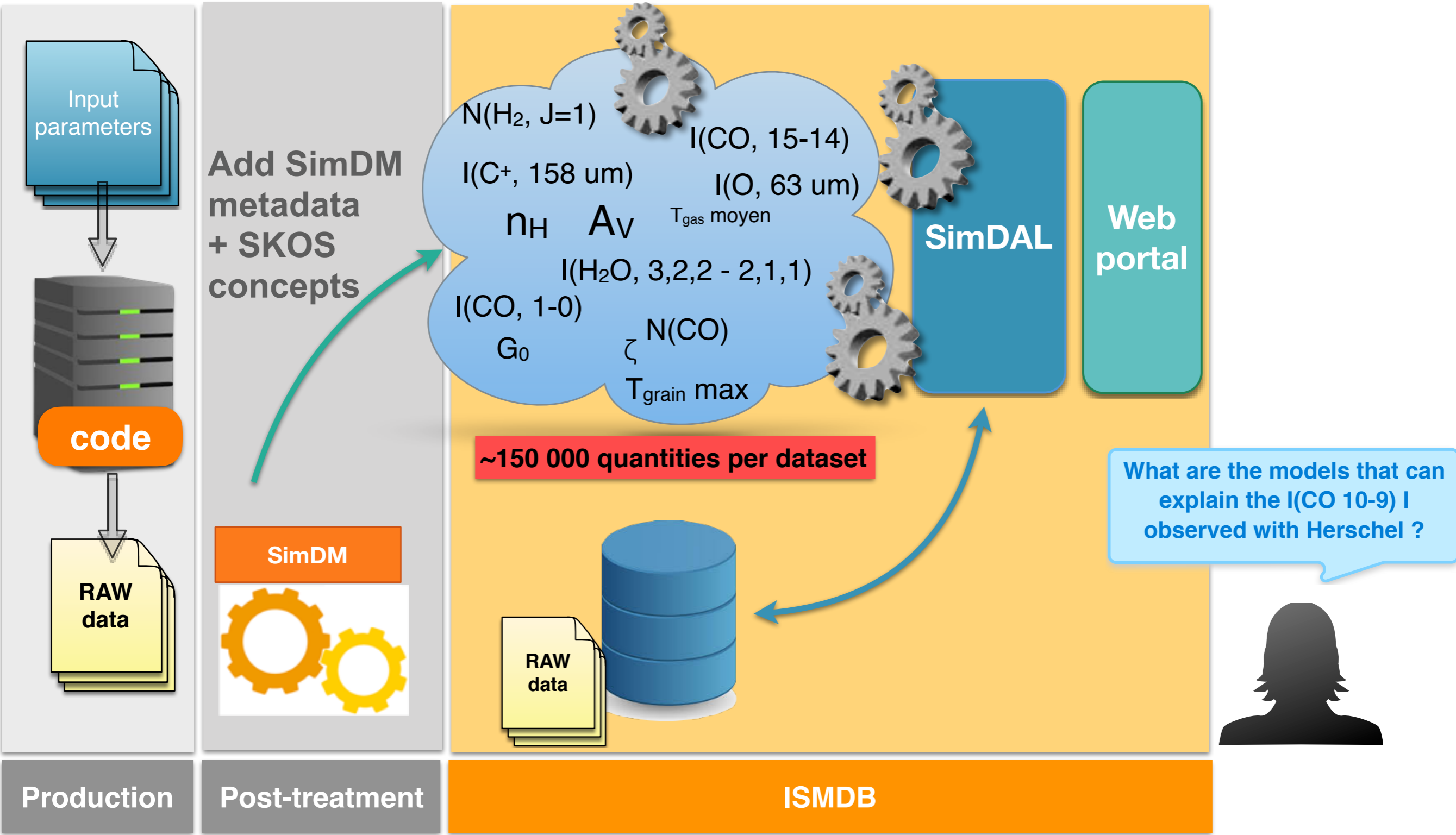


Chevance et al. (2016)



Minimization procedure

Application of vocabularies: ISMDB



Application of vocabularies: ISMDB



CODES

ISMDB

PARTNERS

REGISTRATION

Help

Contact

ISM DataBase – Inverse Search service Beta

Grid of isobaric PDR 1.5.2 models

2016.12.03

1 – search among two parameters

x (cm⁻³_K) log scale
y (Mathis_unit) log scale

2 – fix all the other parameters

(mag)

3 – observational constraints

Search for available quantities... Ex: N(H)

Use

```
"I(CO v=0J=1->v=0J=0 angle 00 deg)" > 1.8E-7  
"I(CO v=0J=1->v=0J=0 angle 00 deg)" < 2.4E-7  
"I(H2 v=0J=2->v=0J=0 angle 60 deg)" > 1E-8  
"I(H2 v=0J=2->v=0J=0 angle 60 deg)" < 5E-7
```

Search

① Select the searched input parameters

Example of a search:

- gas pressure
- UV intensity

② Fix the other input parameters

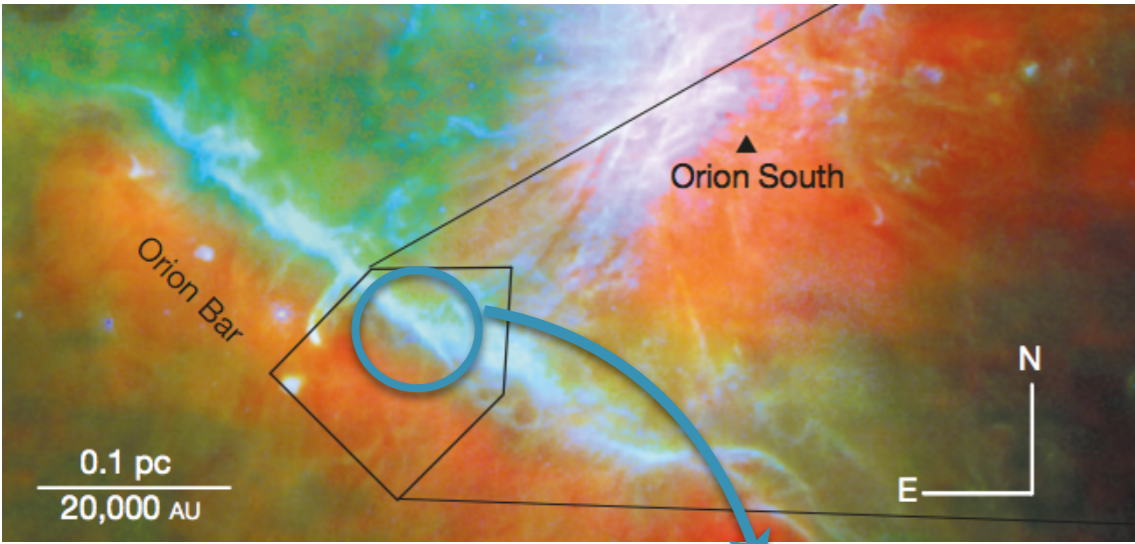
Example: size of the cloud

③ Enter the observations

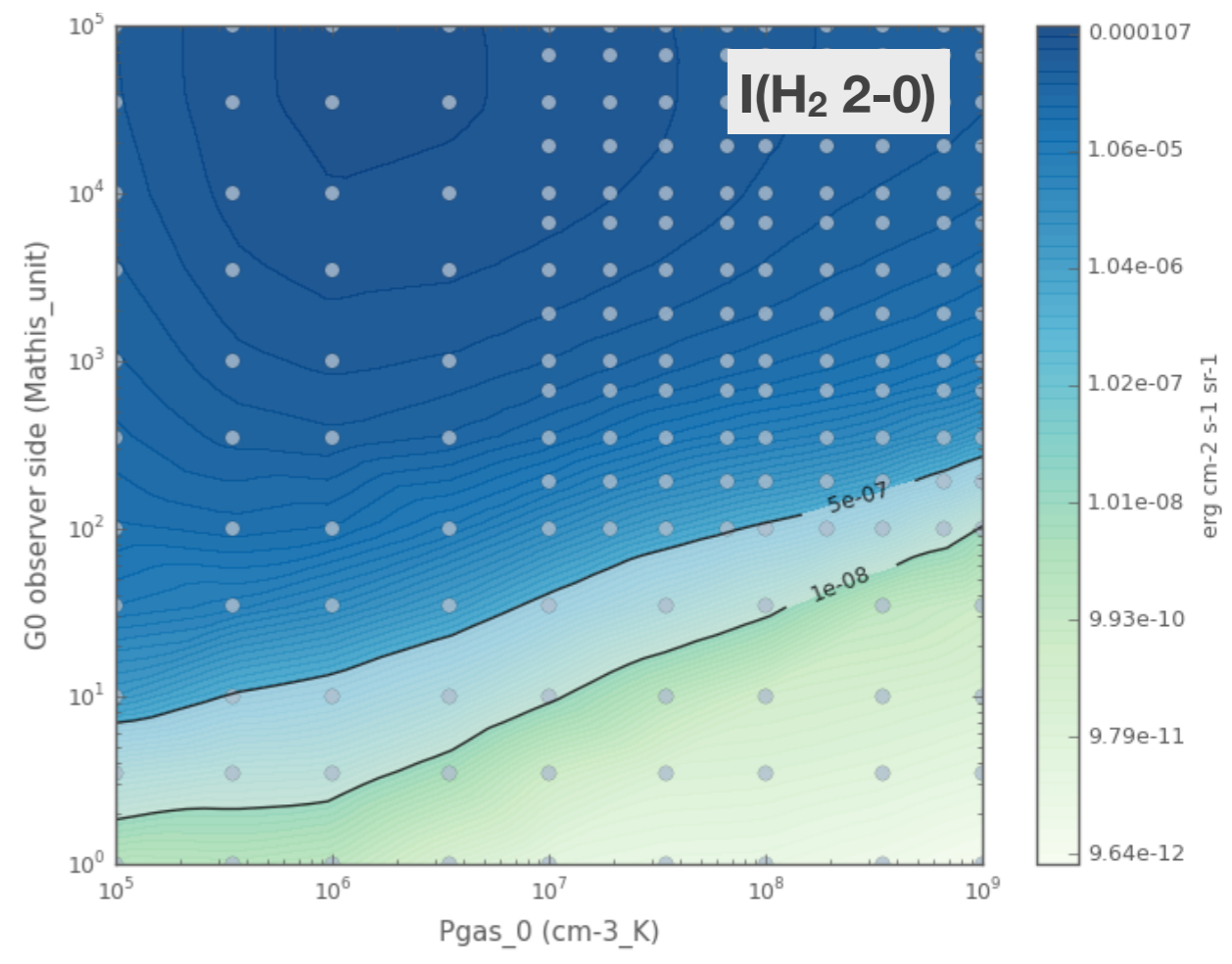
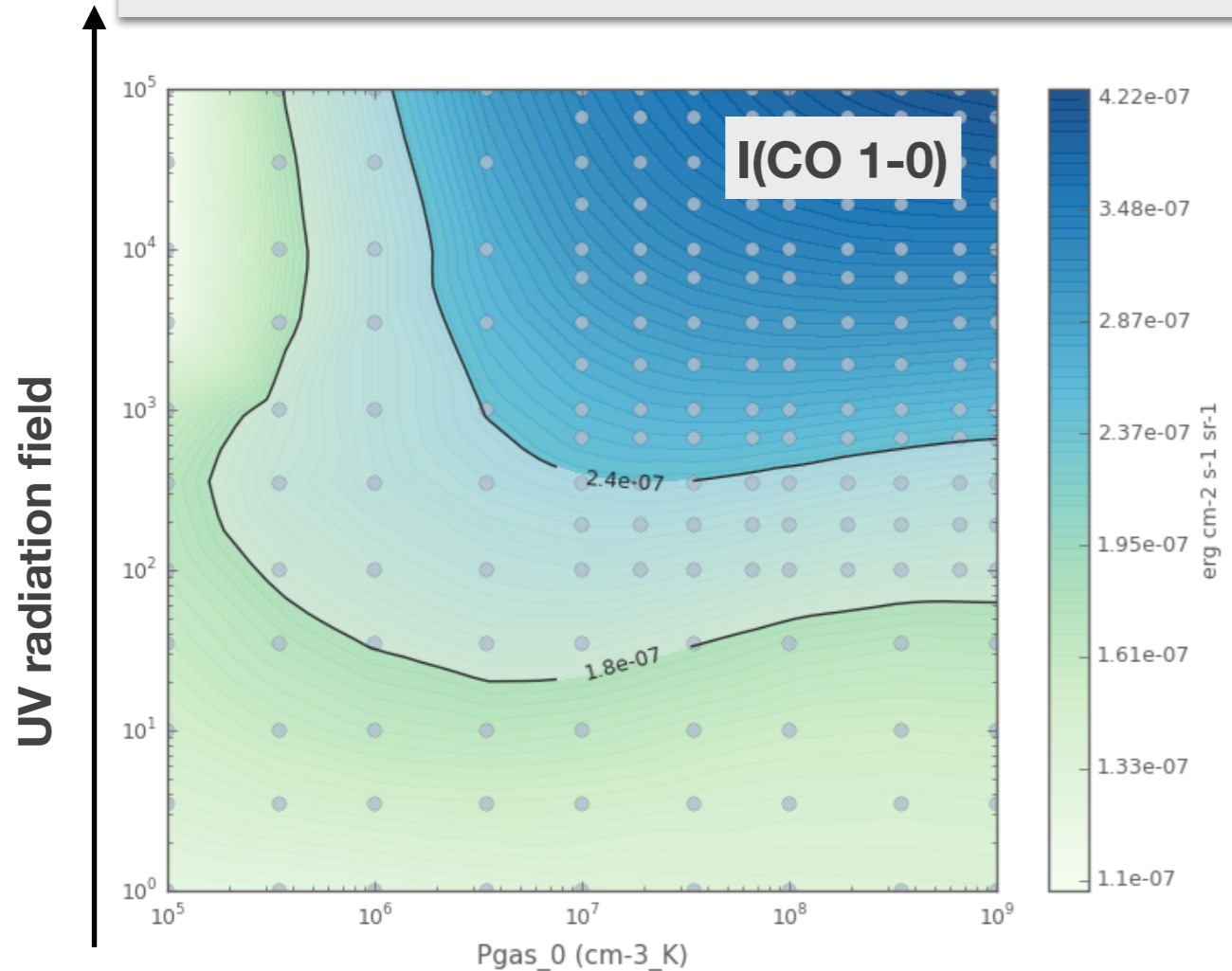
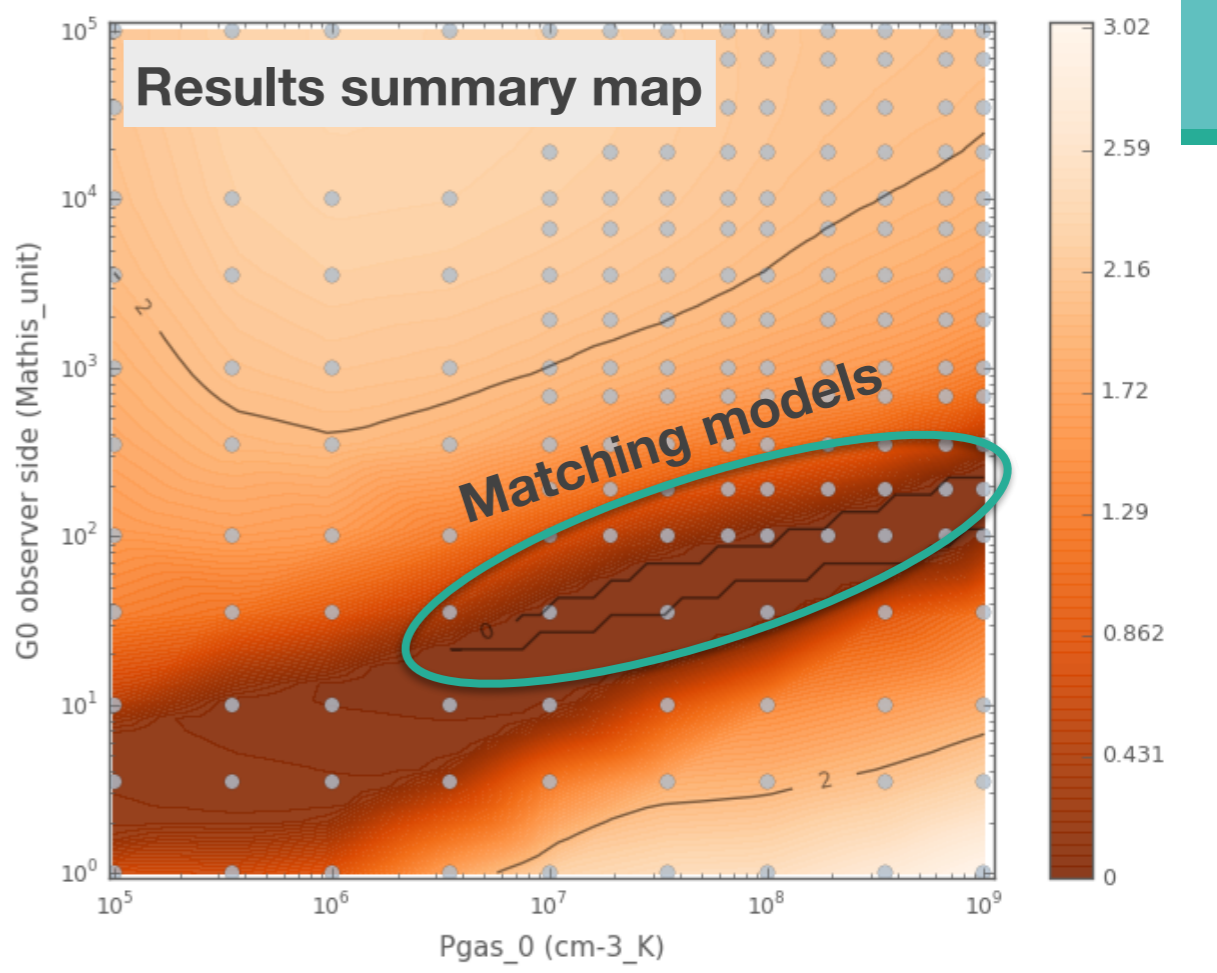
Example: observations CO and H₂ intensities

$$1.8 \cdot 10^{-7} < I(\text{CO } 1-0) < 2.4 \cdot 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$$
$$1.0 \cdot 10^{-8} < I(\text{H}_2 \text{ } 2-0) < 5.0 \cdot 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$$

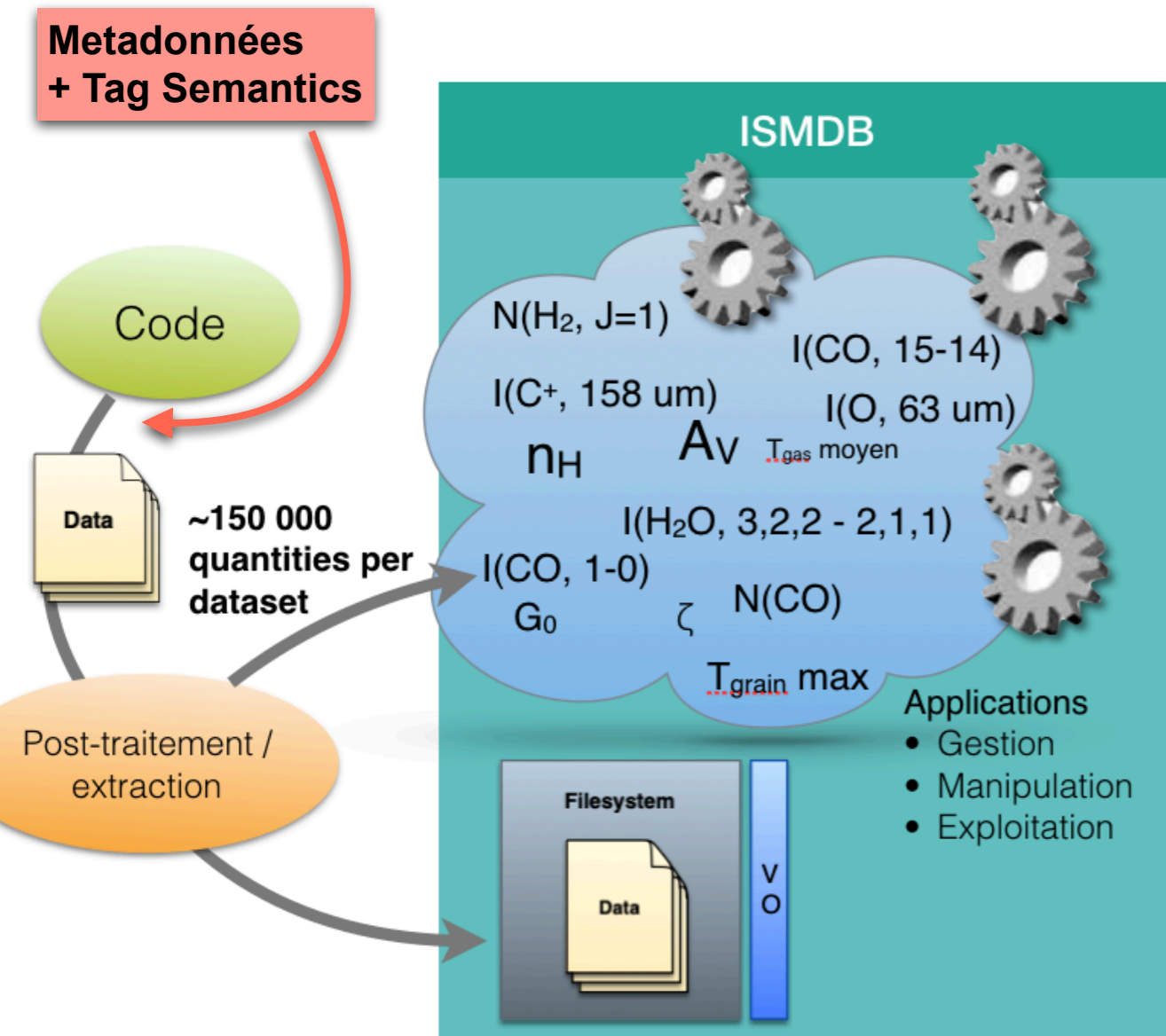
Application of vocabularies: IS



$1.8 \cdot 10^{-7} < I(\text{CO } 1-0) < 2.4 \cdot 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$
 $1.0 \cdot 10^{-8} < I(\text{H}_2 \text{ } 2-0) < 5.0 \cdot 10^{-7} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$



Application of vocabularies: ISMDB

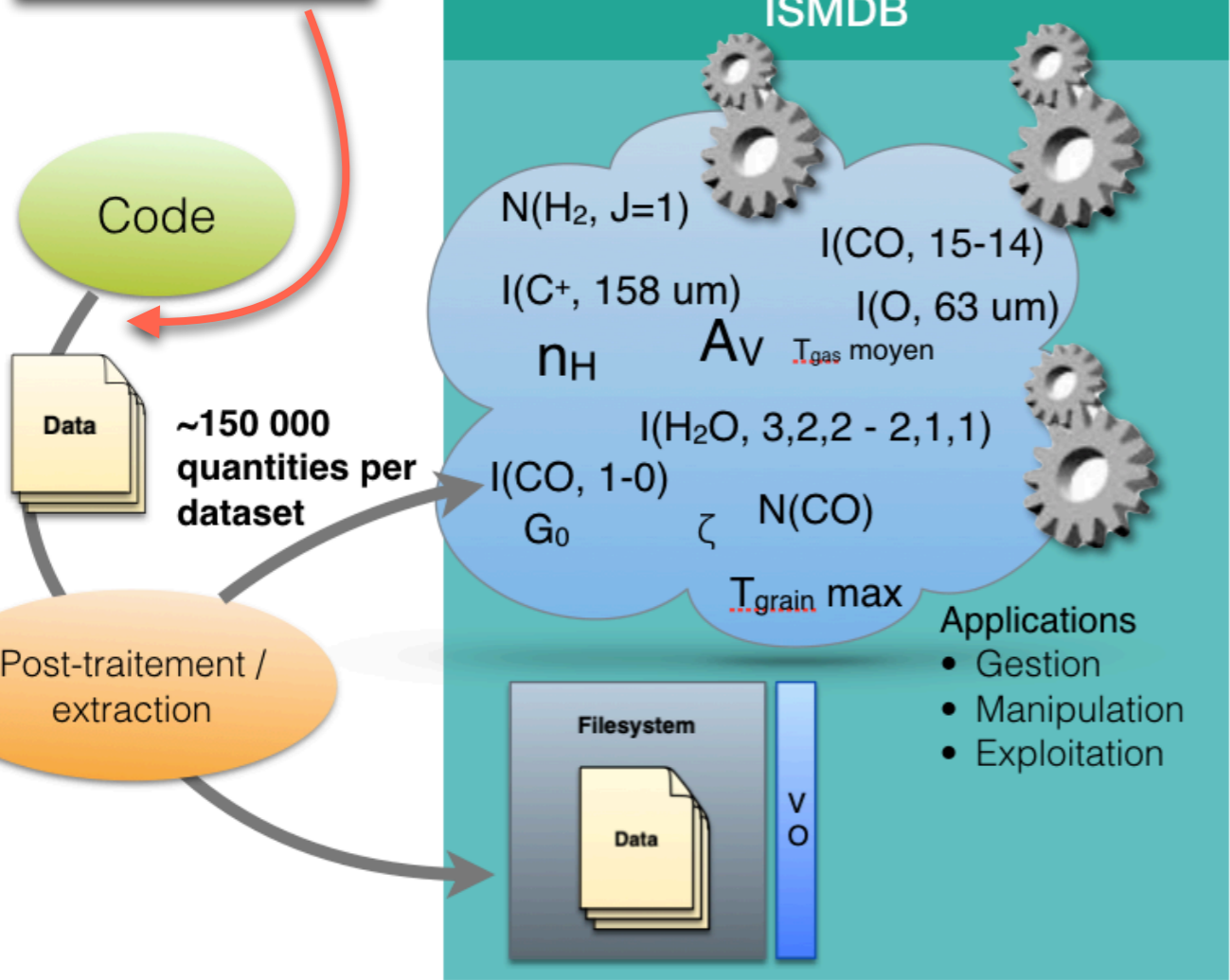


→ **High dimension** database

→ Interrogations possible on more than 150 000 quantities

Application of vocabularies: ISMDB

Metadonnées
+ Tag Semantics



Example of a Classical interface (ex: VLA)

NRAO Science Data Archive : Advanced Search Tool
Historical VLA, Jansky VLA, VLBA and GBT Data Products

Submit Query Check Query Clear Form

Output Control Parameters :

Choose Query Return Type :
 Download Archive Data Files
 VLA Observations Summary
 List of Observation Scans
 List of Projects

Output Tbl Format: HTML
 Max Output Tbl Rows: NO LIMIT
 Sort Order Column 1: Starttime Asc
 Sort Order Column 2: Starttime Asc

General Search Parameters :

Telescopes: All Jansky VLA Historical VLA VLBA GBT
 Project Code: GBT: AGBT12A_055 JVA: 12A-256
 Project Session: _____
 Dates From: _____ To: _____ (2010-06-21 14:20:30)
 Observer Name: _____ Archive File ID: _____ (partial strings allowed)

Position Search :

Target Name: _____ Search Type: SIMBAD or NED
 RA or Longitude: (04h33m11.1s or 68.29d) DEC or Latitude: (05d21'15.5" or 5.352d)
 Search Radius: 1.0" (1d00'00" or 0.2d) Min. Exposure: _____ (secs) Equinox: J2000
 Check for automatic VLA field-of-view, freq. dependent.??

Observing Configurations Search :

Telescope: All A AB BnA B BC CnB
 Config: C CD DnC D DA
 Sub_array: All 1 2 3 4 5
 Polarization: ALL
 Data Type: ALL
 Observing Bands: All 4 P L S C X U K Ka Q W
 Frequency Range: _____ (In MHz : 1665.401 - 1720.500)

Enter Locked Project Access key : _____ Unique keywords may be used to unlock proprietary data from individual observing projects. Contact the [NRAO Data Analysts](#) for project access keys.

Submit Query Check Query Clear Form

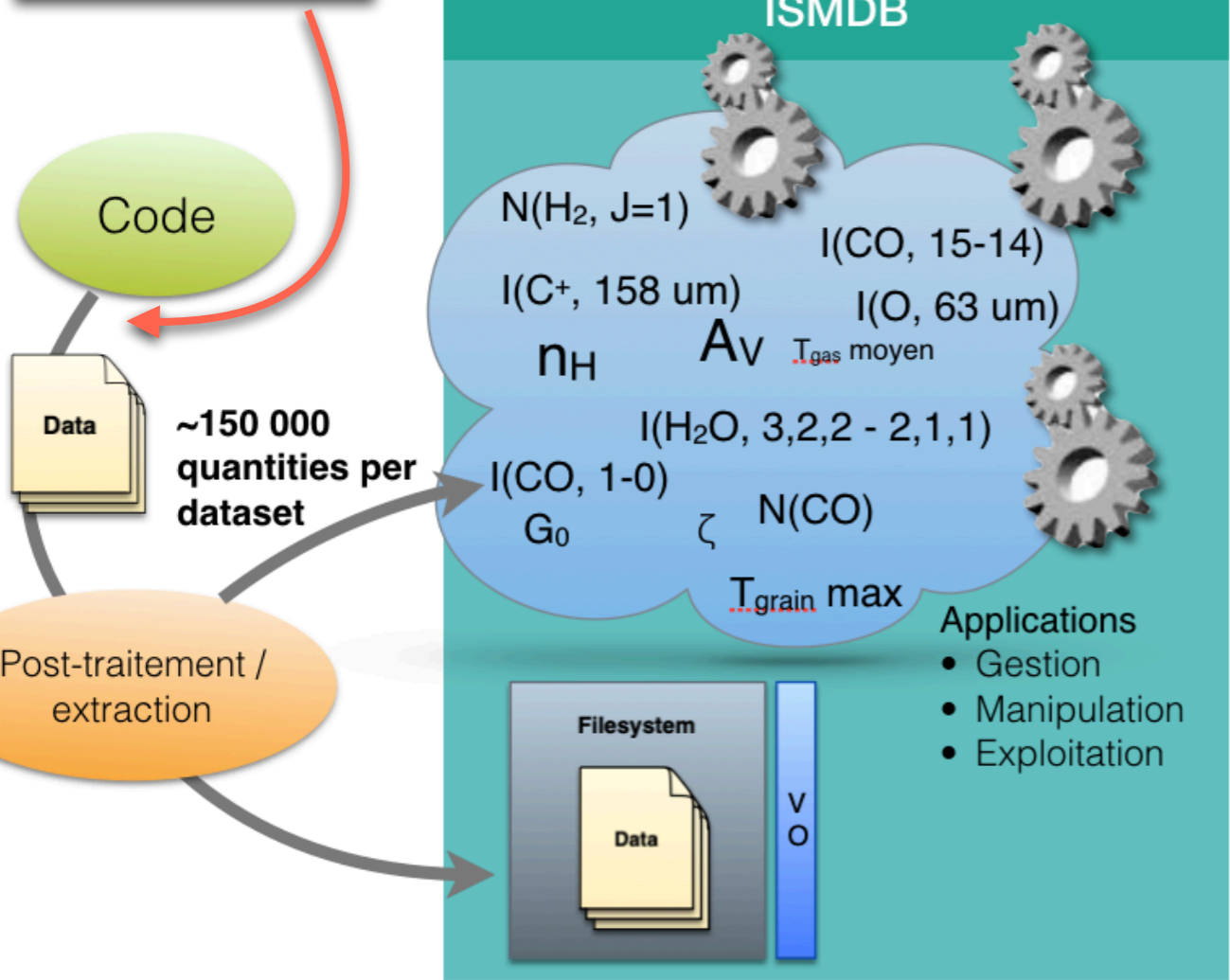
→ High dimension database

→ Interrogations possible on more than 150 000 quantities

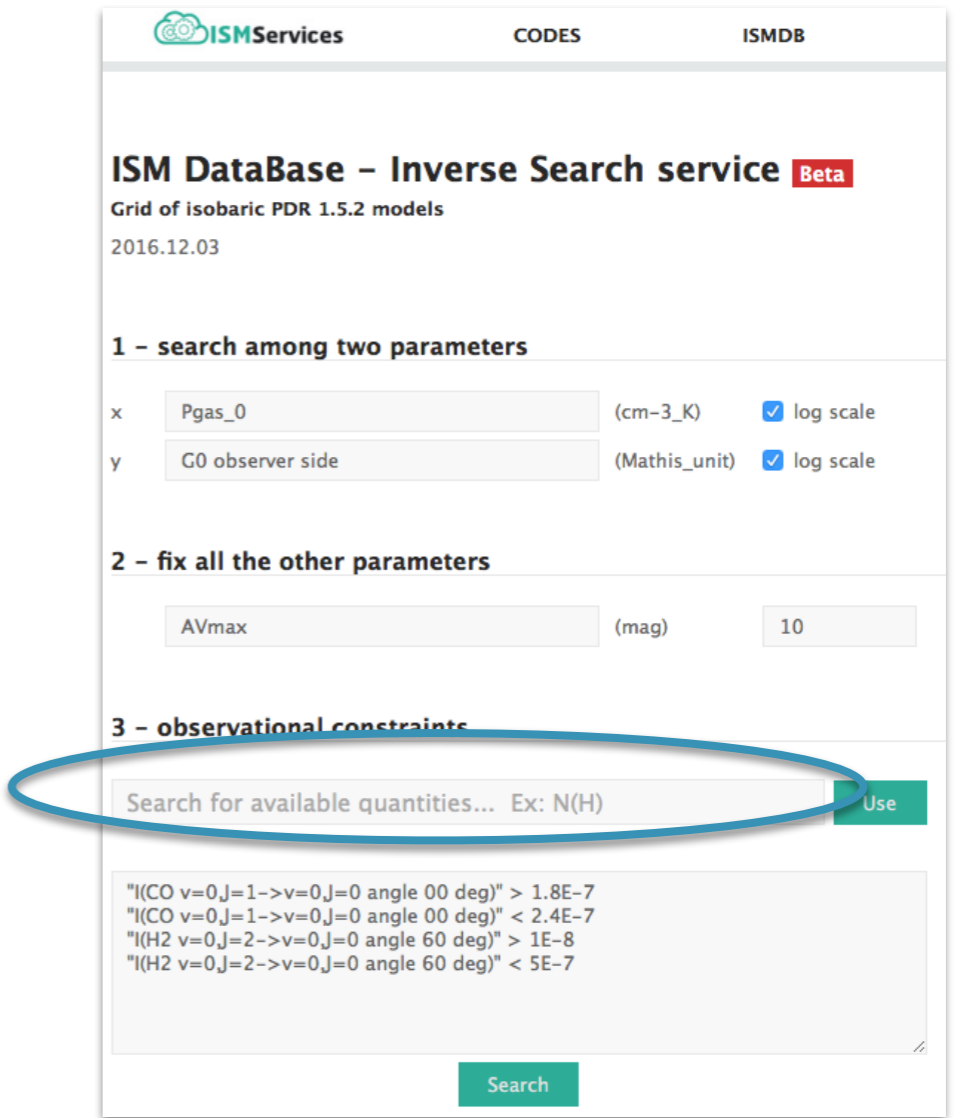
For ISMDB, ~ 150 000 fields would be necessary !

Application of vocabularies: ISMDB

Metadonnées
+ Tag Semantics

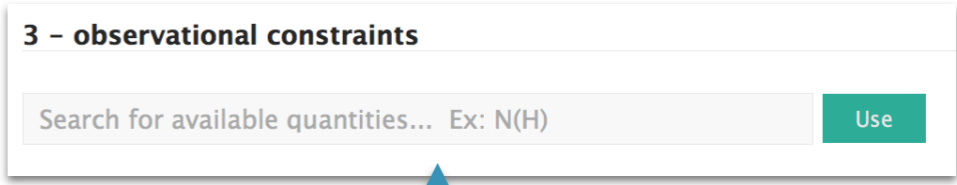


ISMDB interface



→ High dimension database

→ Interrogations possible on more than 150 000 quantities

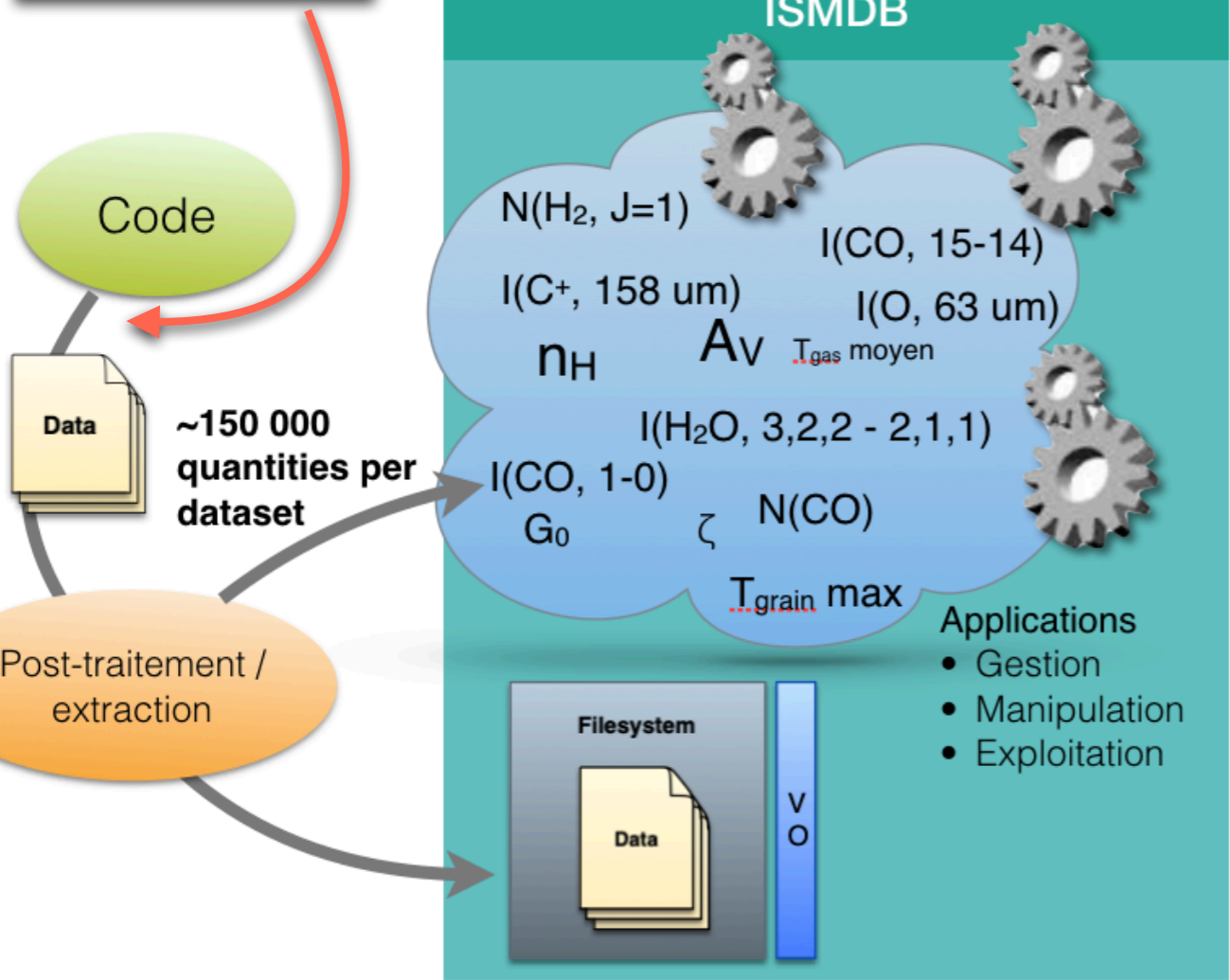


Semantic interpreter

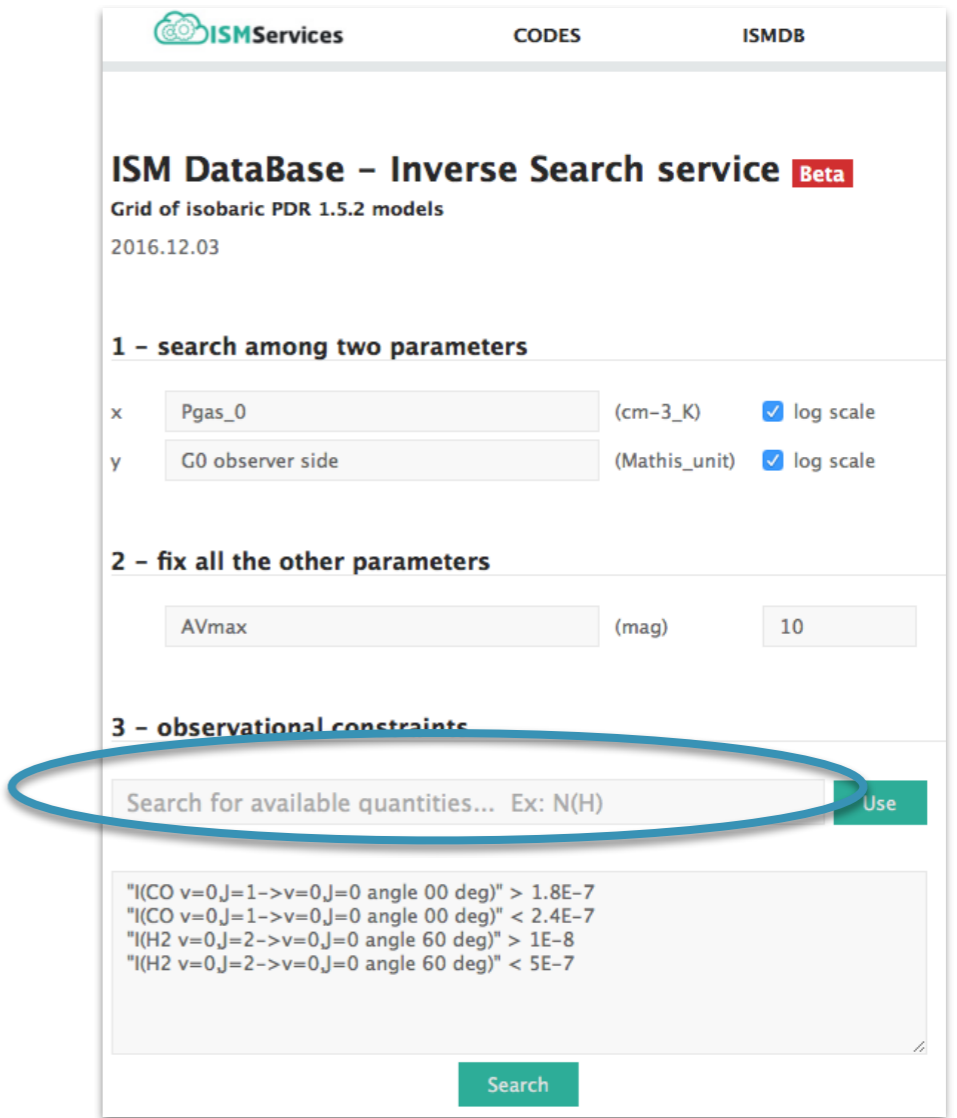


Application of vocabularies: ISMDB

Metadonnées
+ Tag Semantics



ISMDB interface



→ High dimension database

→ Interrogations possible on more than 150 000 quantities

tag each quantity with semantics concept

Semantic interpreter

Application of vocabularies: ISMDB

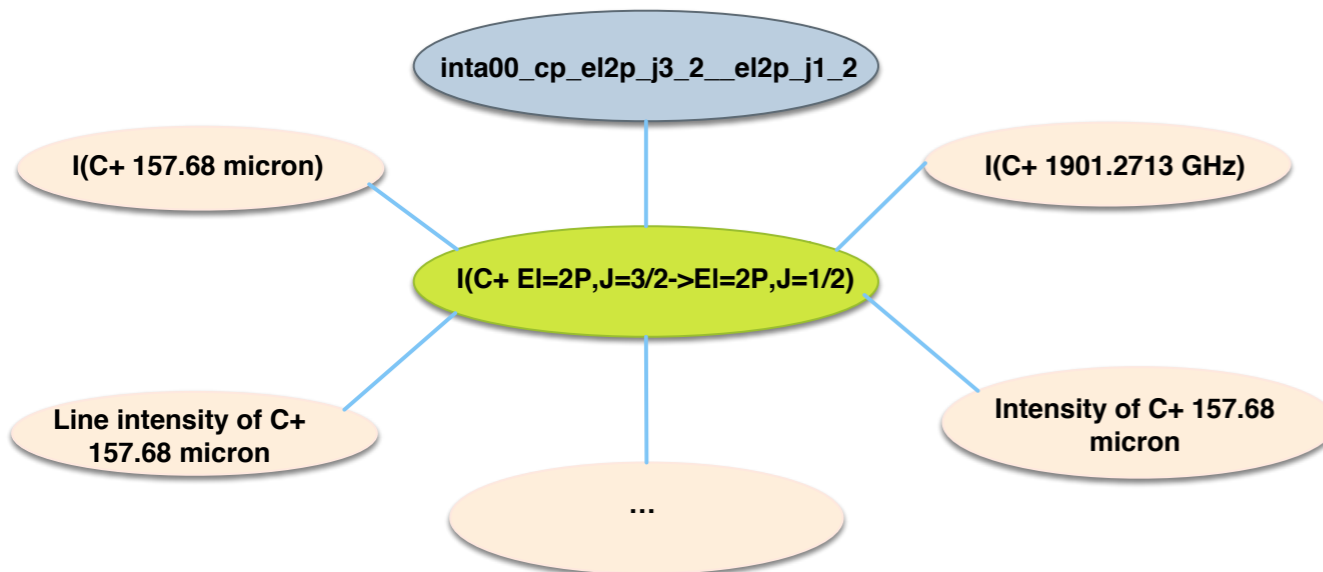
Specific vocabulary used in ISMDB

Exemples:

- ionisation rate by cosmic rays
- mean UV radiation field, G_0
- Line intensity of the H_2^{18}O between levels 1_{10} and 0_{00}
- colonne de densité de CO dans son état $J = 1$
- ...

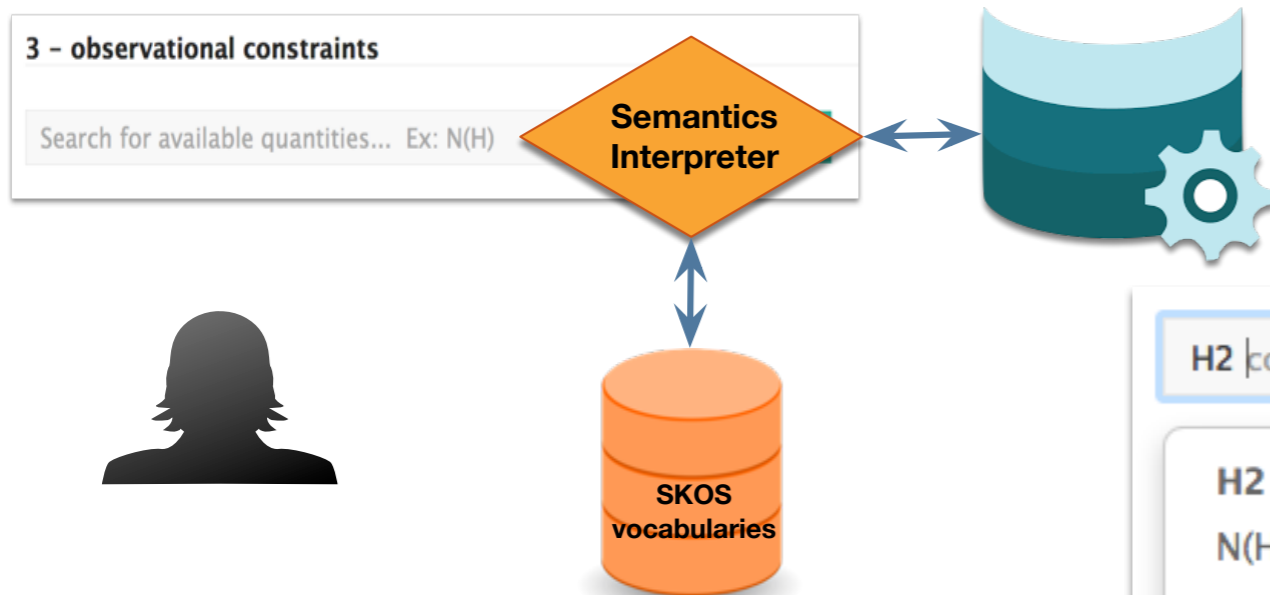
→ specific vocabulary of
~ **300 000 concepts** with
synonyms

Example of the 157.7 micron C+ line intensity



ID: inta00_cp_el2p_j3_2__el2p_j1_2
PREF: I(C+ El=2P,J=3/2->El=2P,J=1/2)
ALT: I(C+ El=2P,J=3/2->El=2P,J=1/2) face on
ALT: I(C+ 157.68 micron) face on
ALT: Intensity of C+ 157.68 micron face on
ALT: Line intensity of C+ 157.68 micron face on
ALT: I(C+ 1901.2713 GHz) face on
ALT: Intensity of C+ 1901.2713 GHz face on
ALT: Line intensity of C+ 1901.2713 GHz face on
...

Application of vocabularies: ISMDB



Semantics Interpreter

Semantics

SKOS: PREF + ALT
→ synonyms

+

Ranking system

(learn from users)

The screenshot shows the ISMDB interface with two search results. Each result has a search bar, a list of suggestions, and a 'Use' button.

Search 1: H2 |column density

- H2 column density
- N(H2)
- N(C2H2)
- N(c-C3H2)
- N(C_13CH2)
- N(C_13CH2+)
- C2H2 column density
- Column density of H2

Search 2: I(H2 |p=0 S(0)) angle 00 degrees

- I(H2 0-0 S(0)) angle 00 degrees
- I(H2 10-10 S(0)) angle 00 degrees
- I(H2 9.6645 micrometres) angle 00 degrees
- I(H2 28.2196 micrometres) angle 00 degrees
- I(H2 156.4883 micrometres) angle 00 degrees
- I(H2 v=0,J=2->v=0,J=0) angle 00 degrees
- I(H2O 6.1140 cm⁻¹) angle 00 degrees
- I(H2O J=1,ka=1,kc=1->J=0,ka=0,kc=0) angle 00 degrees