Semantics & Theory

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

SimDM & Semantics

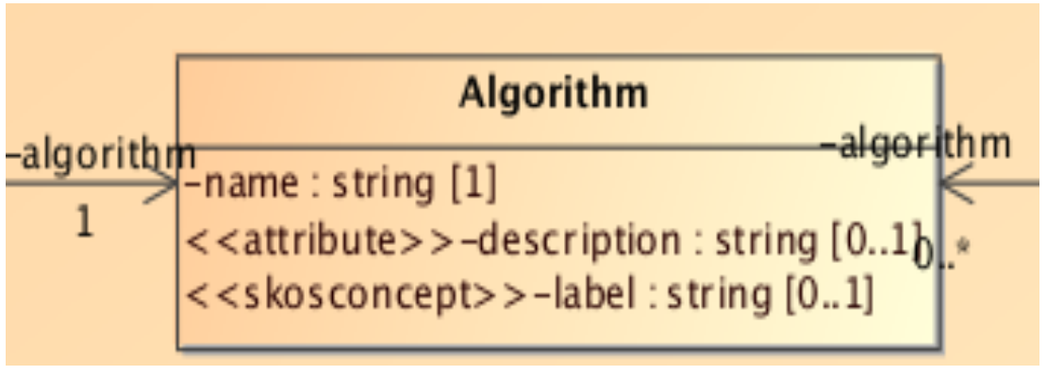
Simulation Data Model (SimDM) is a data model to describe numerical simulations

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 characterise quantities

SimDM + vocabularies

SimDM uses **SKOS** (Simple Knowledge Organization System) vocabularies
- some classes have skosconcept attribute



Theory vocabularies

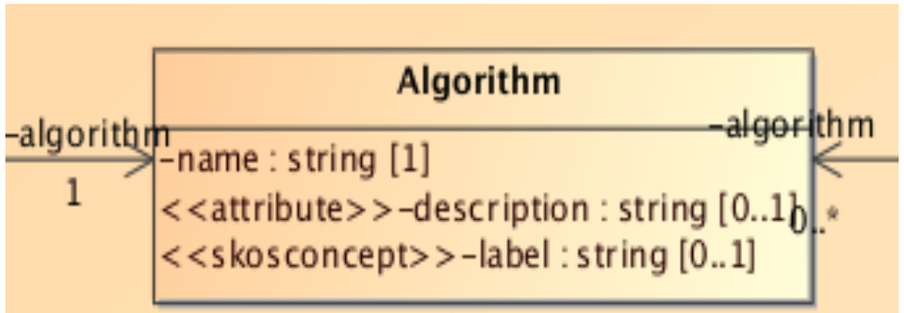
Theory SKOS vocabularies:

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

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

Common name

Synonyms



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 <u>specific to a code</u>
  Example: Choice between home-made algorithms to simulate a process

- some concepts do <u>not have precise definition</u>
  Example: Mean UV radiation field intensity (G_0) : definition depends on the author

- <u>complex</u> 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 **<u>2 kinds of vocabularies</u>**

- **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.

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**

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

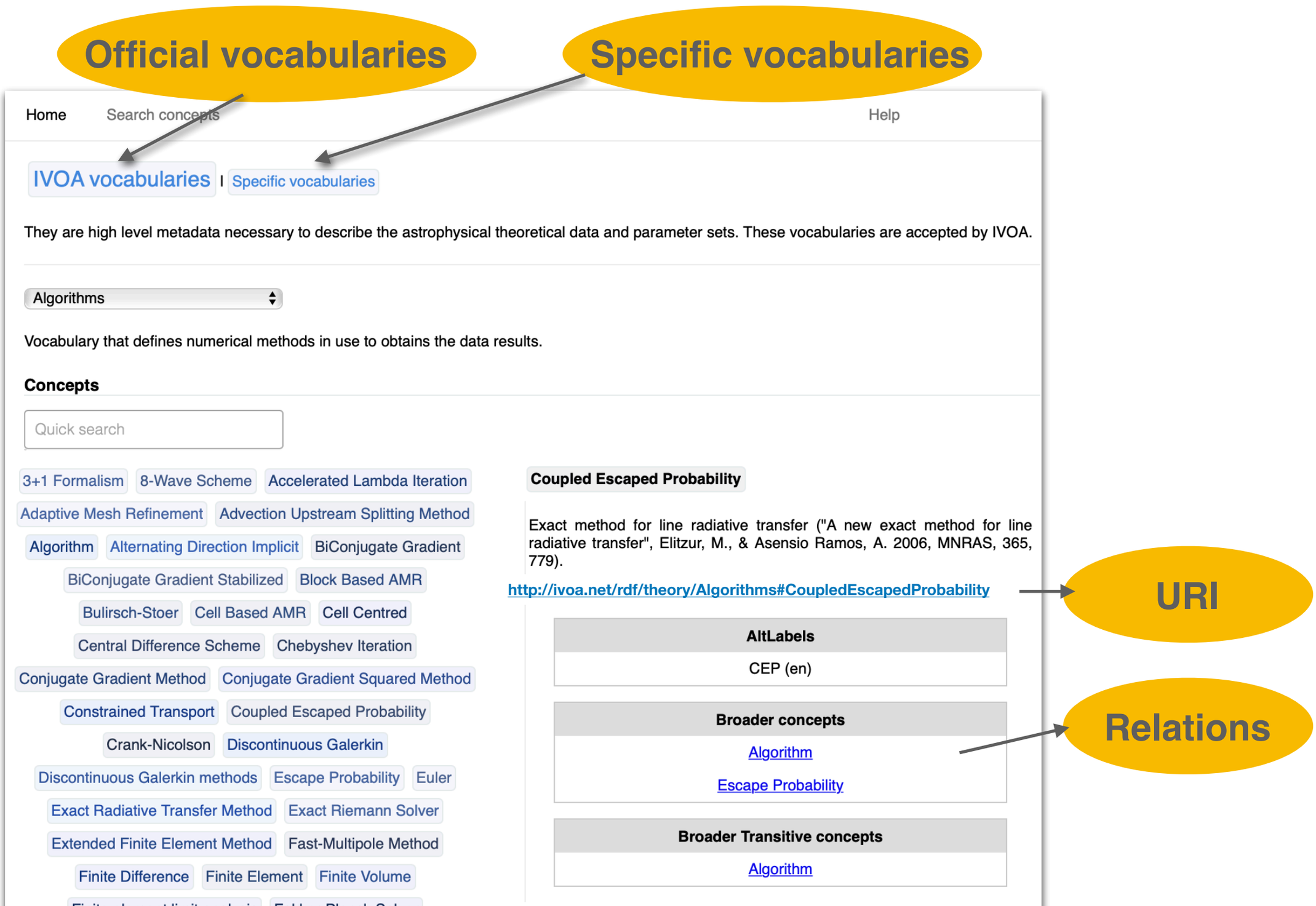
Example: ISMService - http://ism.obspm.fr
Need to describe each atomic and molecular line + many other concepts
➔ specific vocabulary with ~ **300 000 concepts**

Discovering concepts for publication of simulations

Need convenient GUI so scientists can find URIs

- discover concepts
- navigate between concepts

http://votheory.obspm.fr



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 Shangai 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 2018/07/30
  - Store the vocabularies

and the implementation is: 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.votheory@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 re- lies on the slicing of the four-dimensional spacetime by three-dimensional surfaces (hypersur- faces)	#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.

- Lack of 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 ?

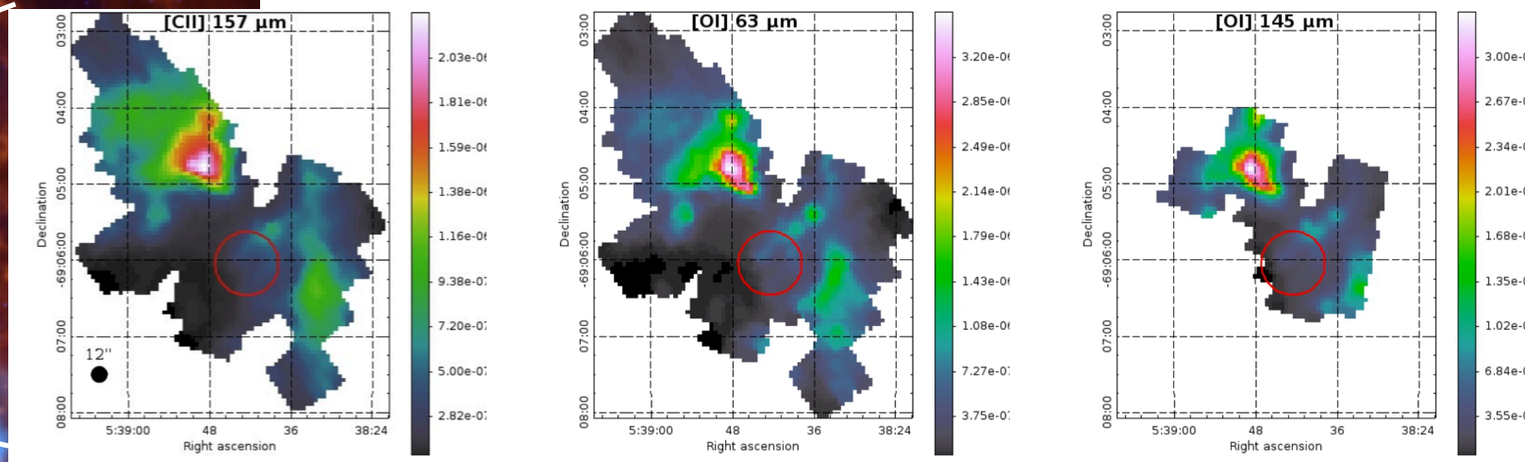
Application of vocabularies: ISMDB

ISMServices

Goal of ISMDB: interpret observations of interstellar gas emission

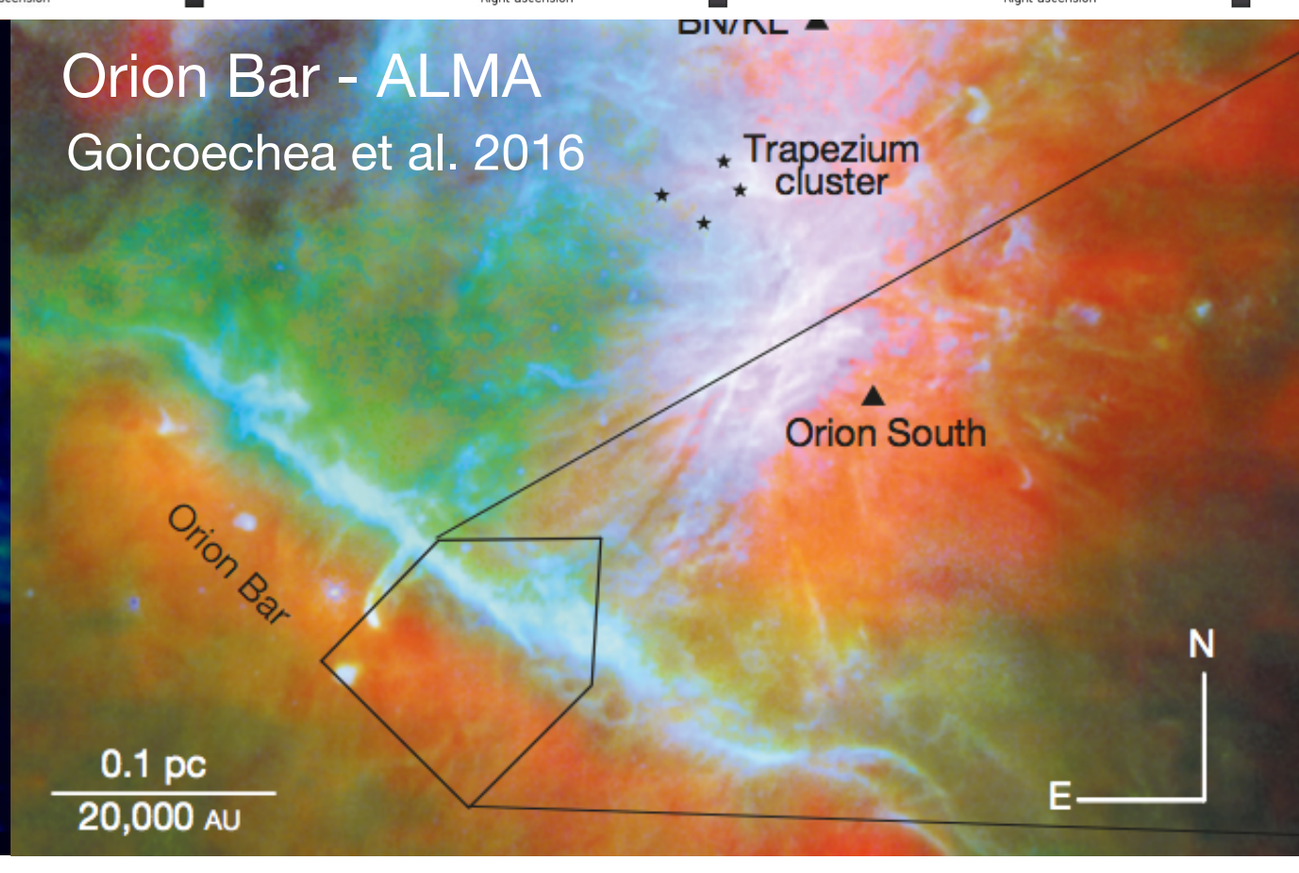
30 Doradus / Large Magellanic Cloud - VISTA/ESO
Chevance et al. 2016

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



Orion B / Horsehead - IRAM 30m
Pety et al. 2017

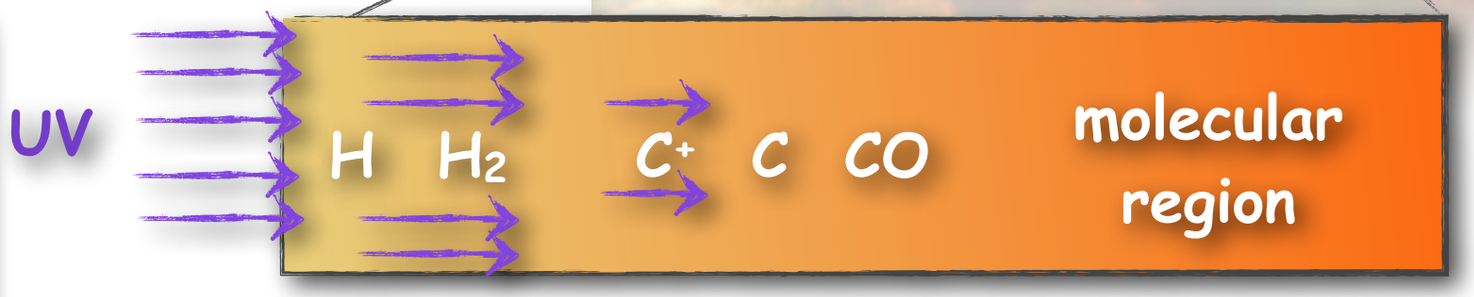
Orion Bar - ALMA
Goicoechea et al. 2016



Application of vocabularies: ISMDB

PDR codes:

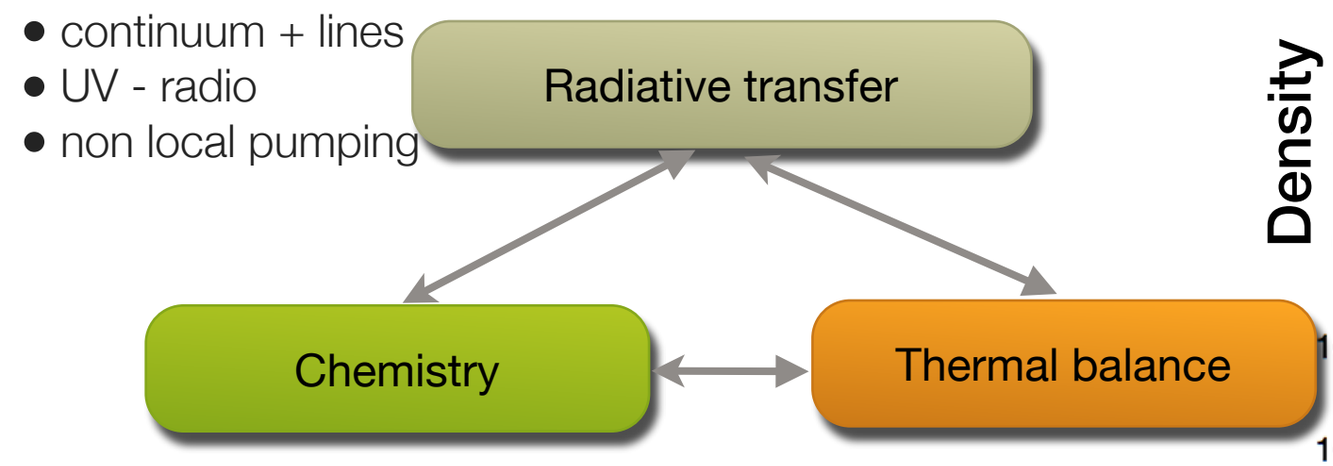
- compute the atomic and molecular structure of interstellar clouds.
- analysis of physical and chemical processes



Outputs

- densities
- excitation
- gas & grains temperatures

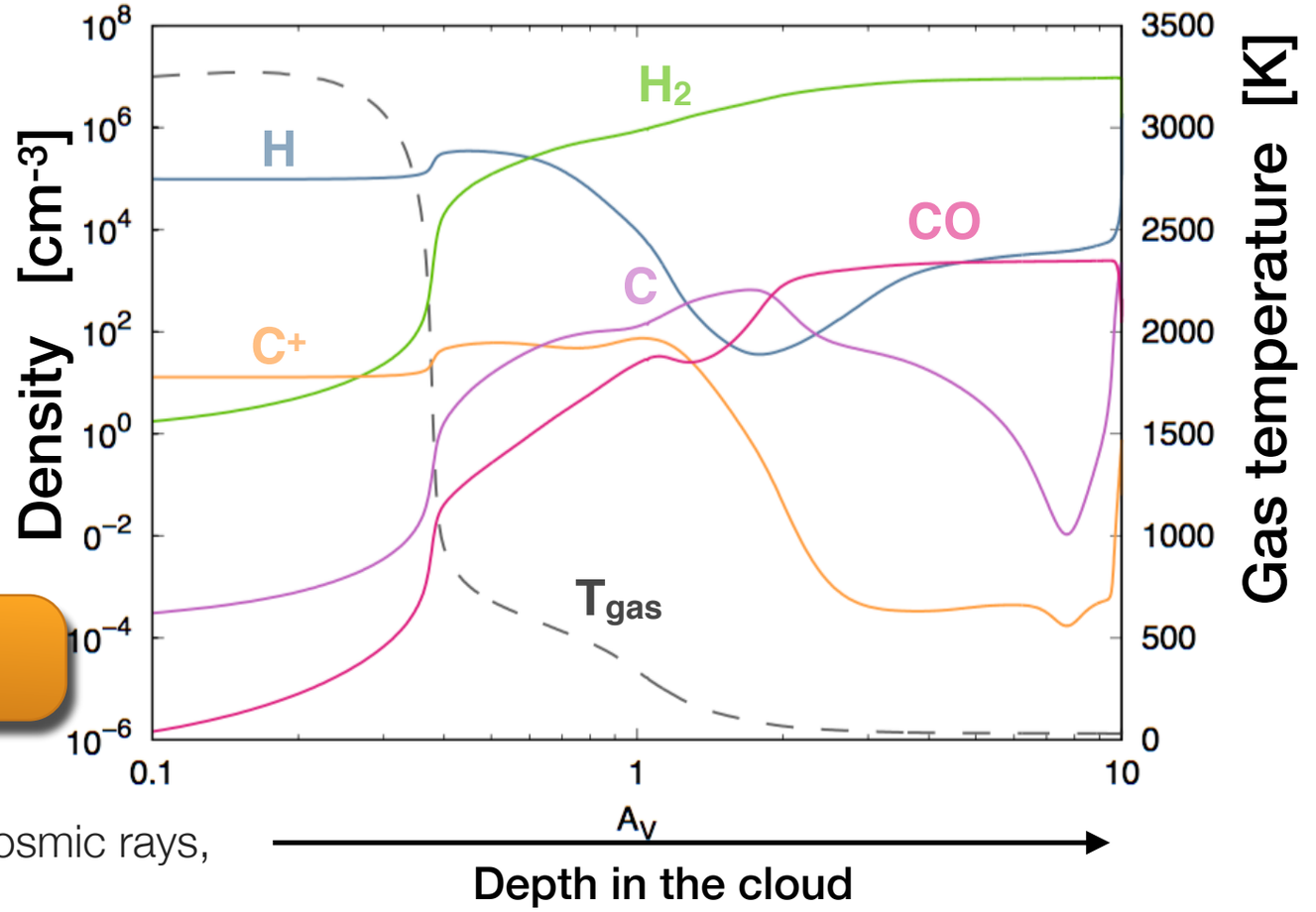
- Intensities (H_2, CO, H_2O, ...)
- Column densities



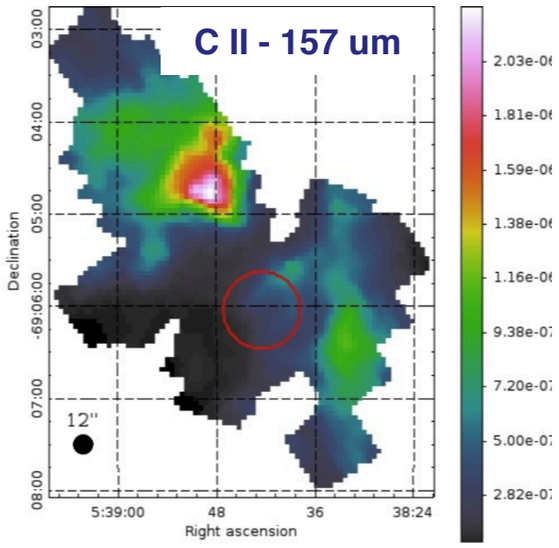
- continuum + lines
- UV - radio
- non local pumping

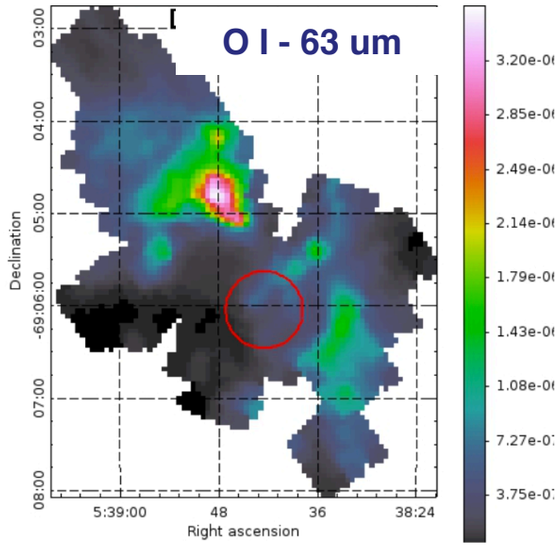
- several hundred species
- thousands reactions
- gas & grains

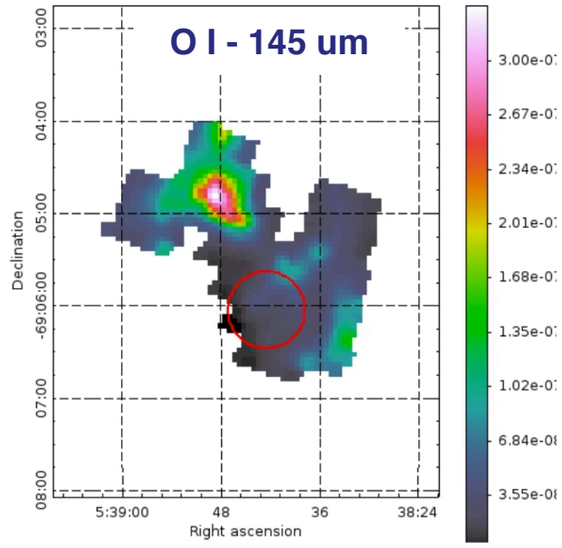
- Statistical equilibrium
- Photo-electric effect, cosmic rays, chemistry, …, ...



Observations







Chevance et al. (2016)

Interpretation

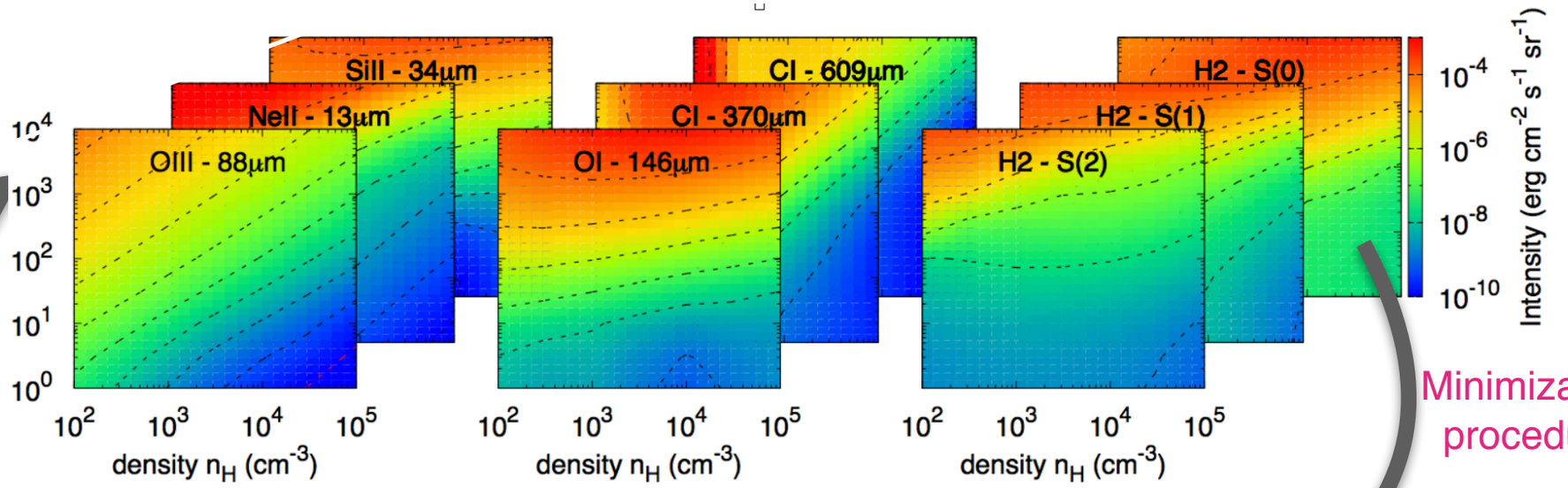
Extract observed line intensities

Database of models

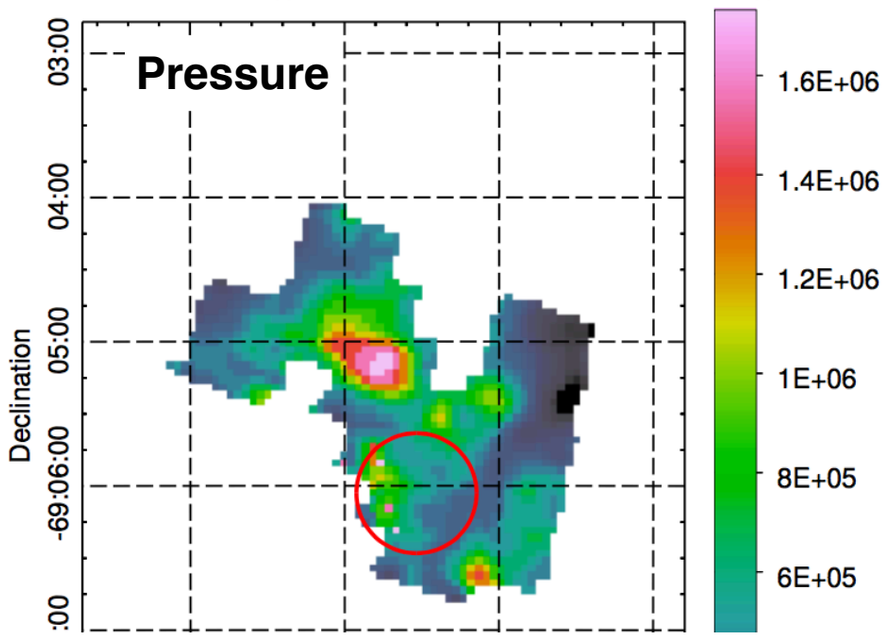
Exploration of parameter space

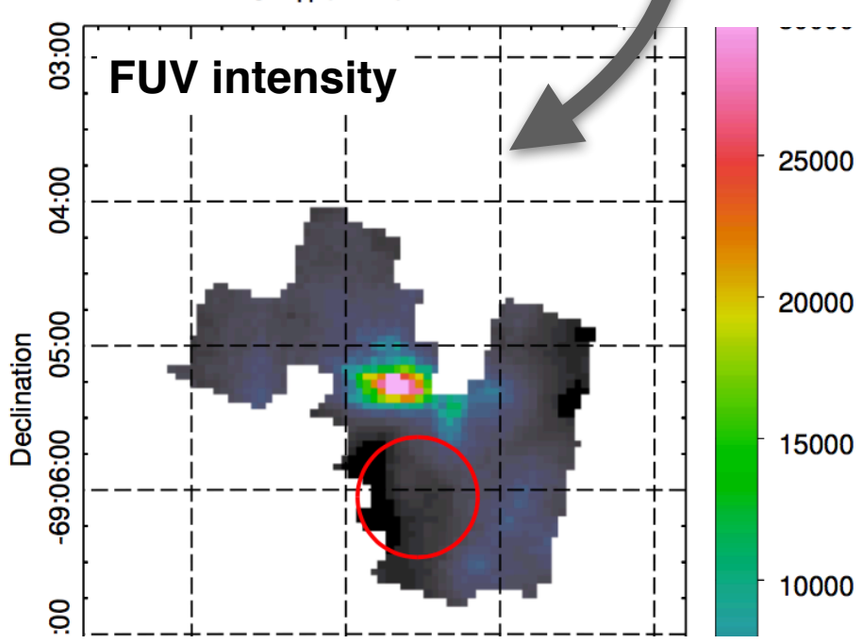
Computes plenty of line intensities

PDR code

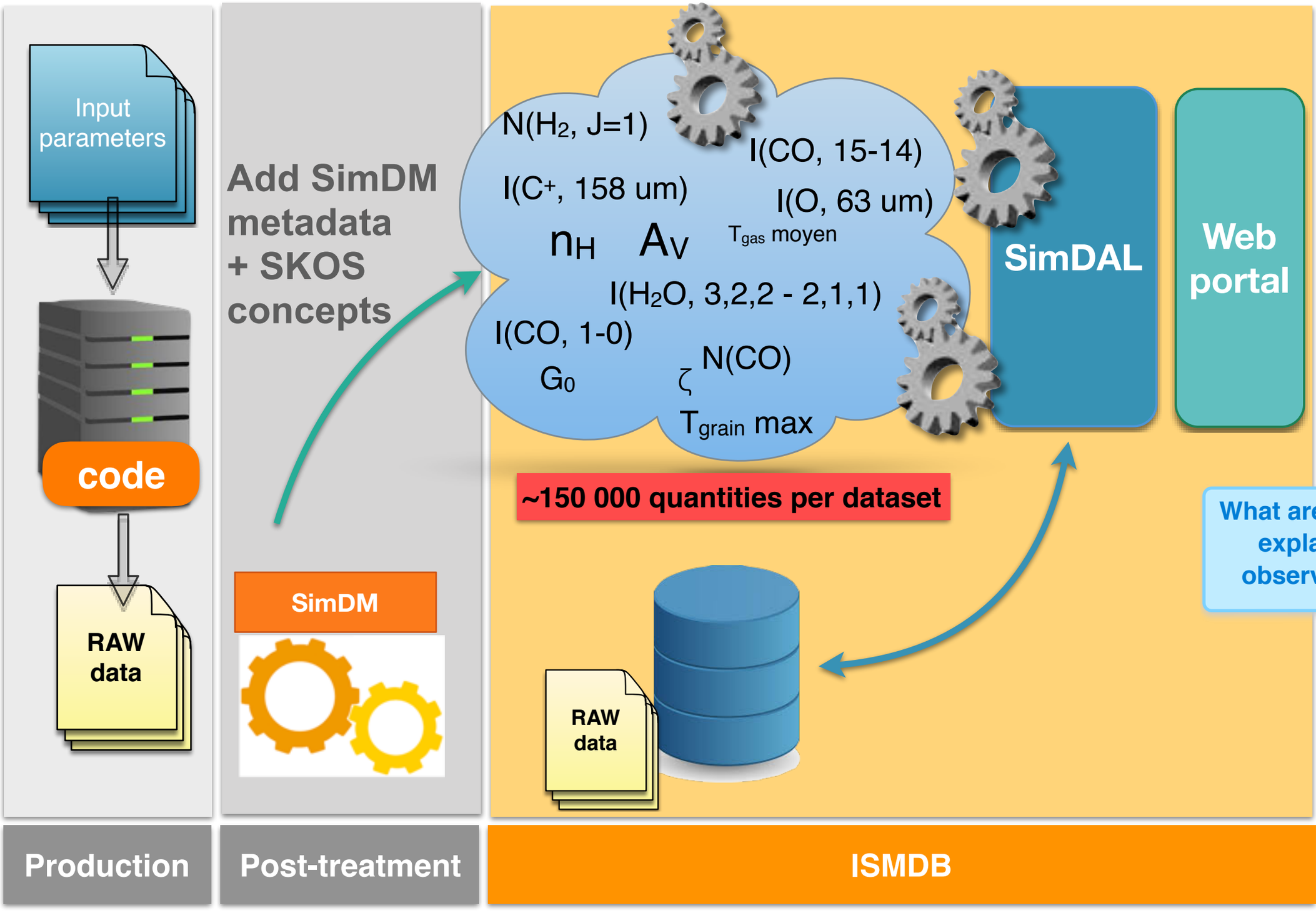


Minimization procedure

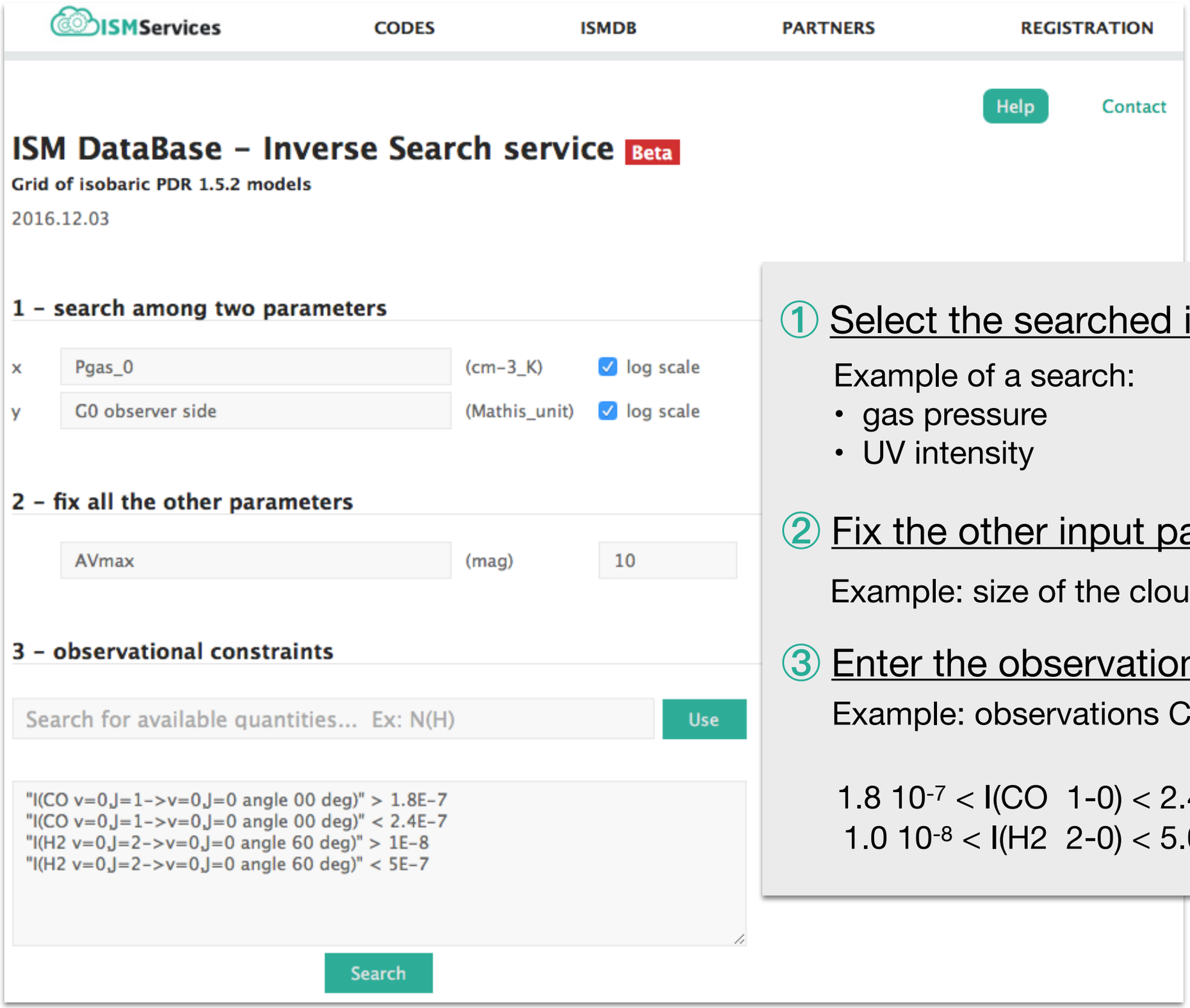




Application of vocabularies: ISMDB



Application of vocabularies: ISMDB



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

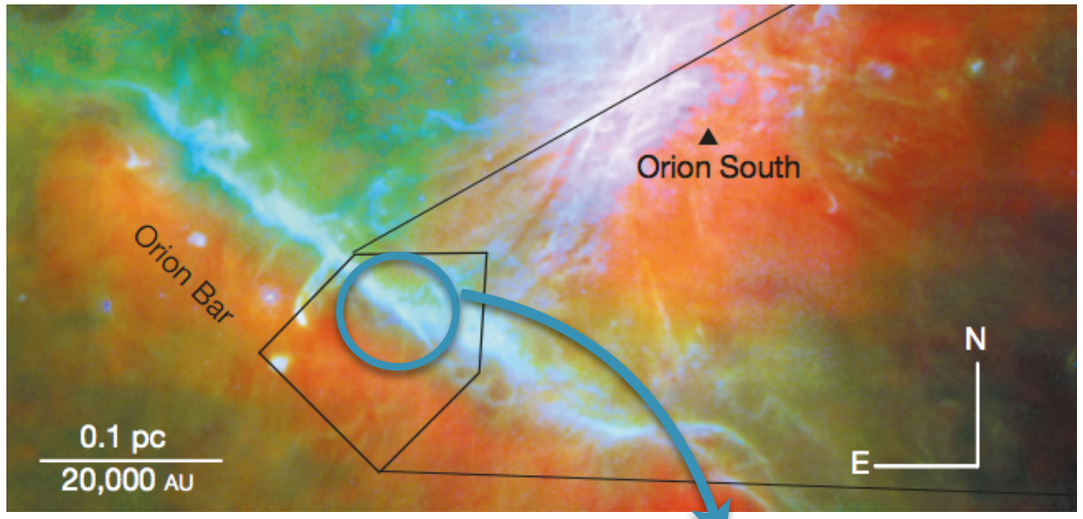
③ Enter the observations

Example: observations CO and H_2 intensities

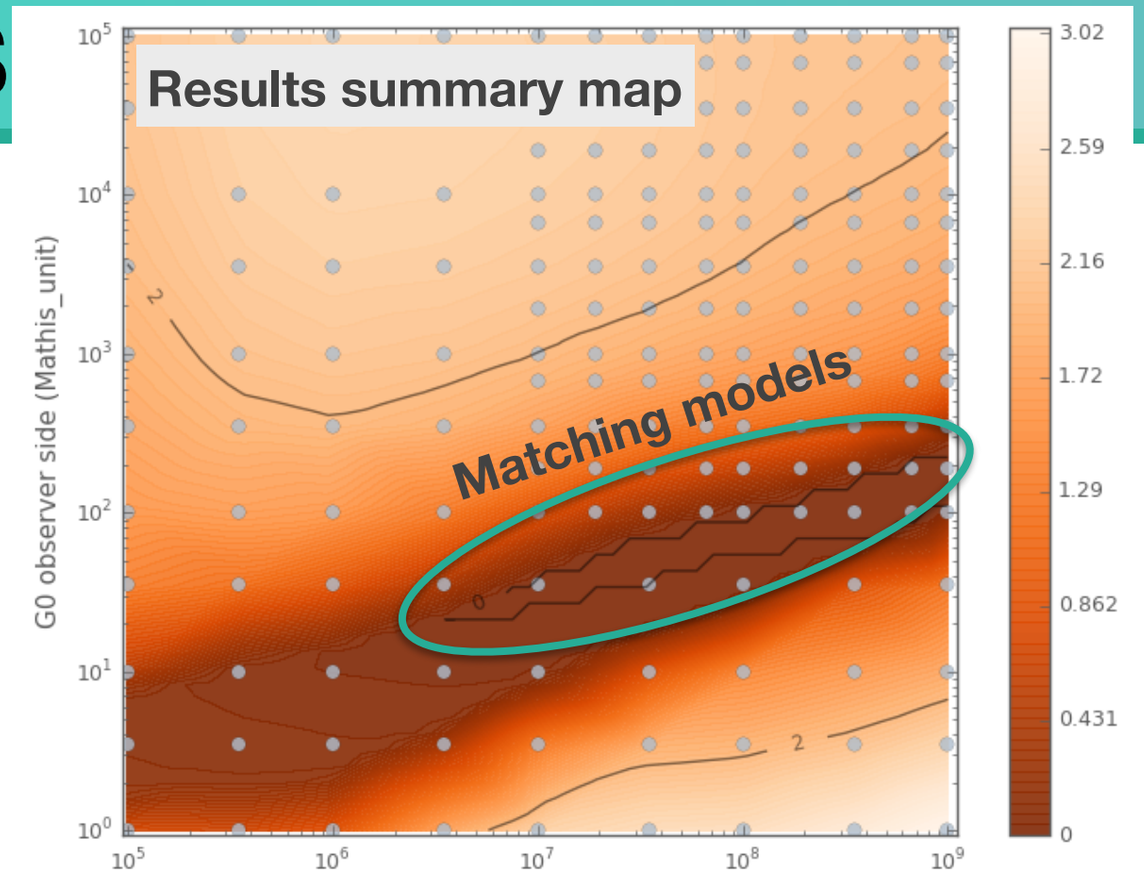
$1.8\ 10^{-7} < I(CO\ \ 1\text{-}0) < 2.4\ 10^{-7}$ erg cm-2 s-1 sr-1

$1.0\ 10^{-8} < I(H2\ \ 2\text{-}0) < 5.0\ 10^{-7}$ erg cm-2 s-1 sr-1

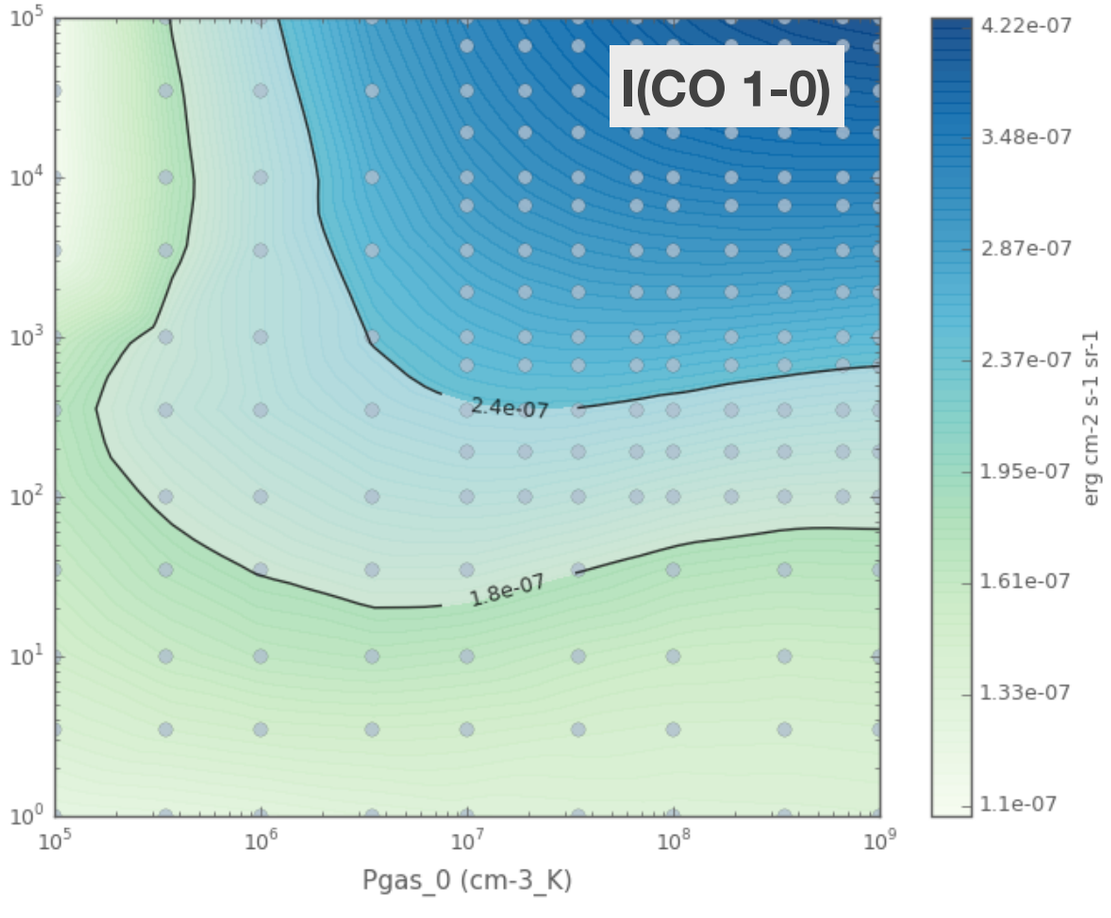
Application of vocabularies: IS

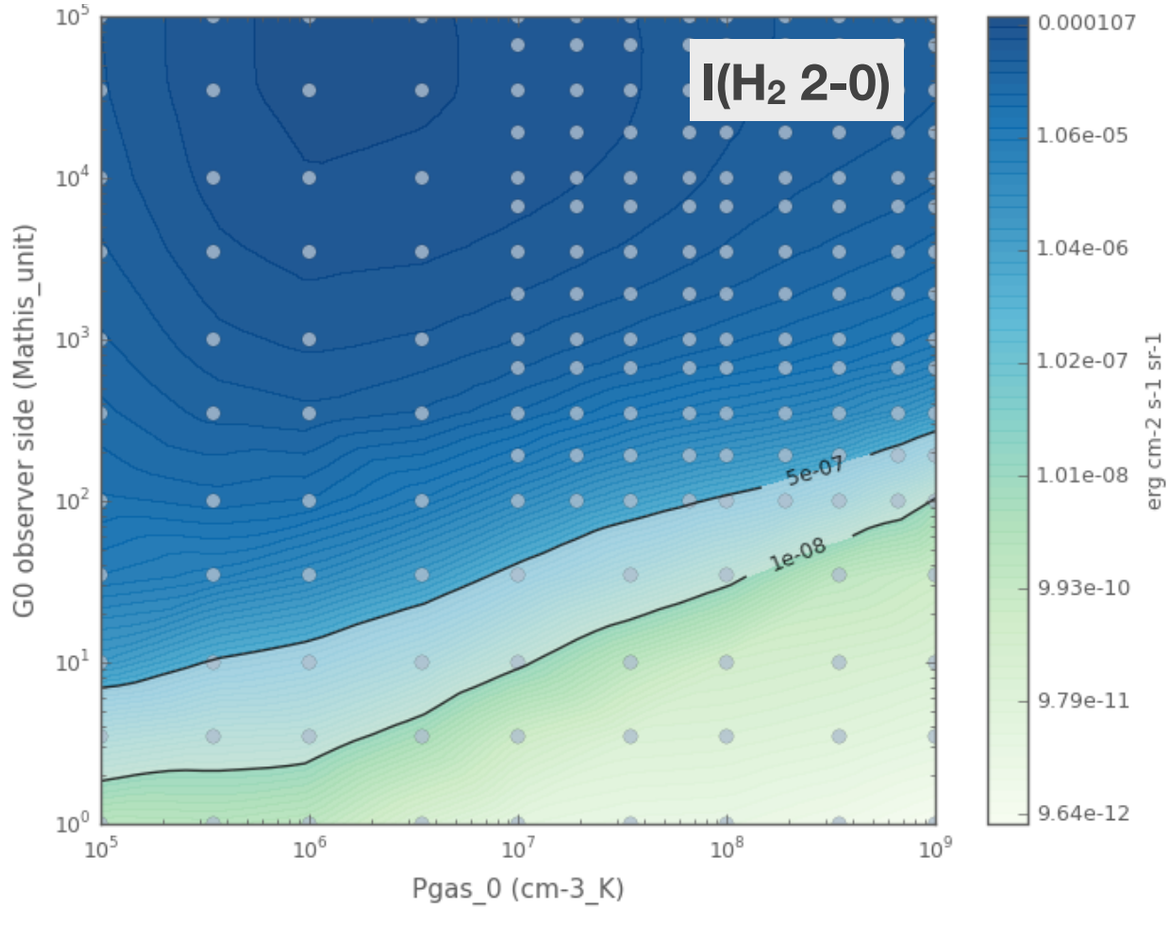


1.8 10^{-7} < I(CO 1-0) < 2.4 10^{-7} erg cm-2 s-1 sr-1
1.0 10^{-8} < I(H2 2-0) < 5.0 10^{-7} erg cm-2 s-1 sr-1

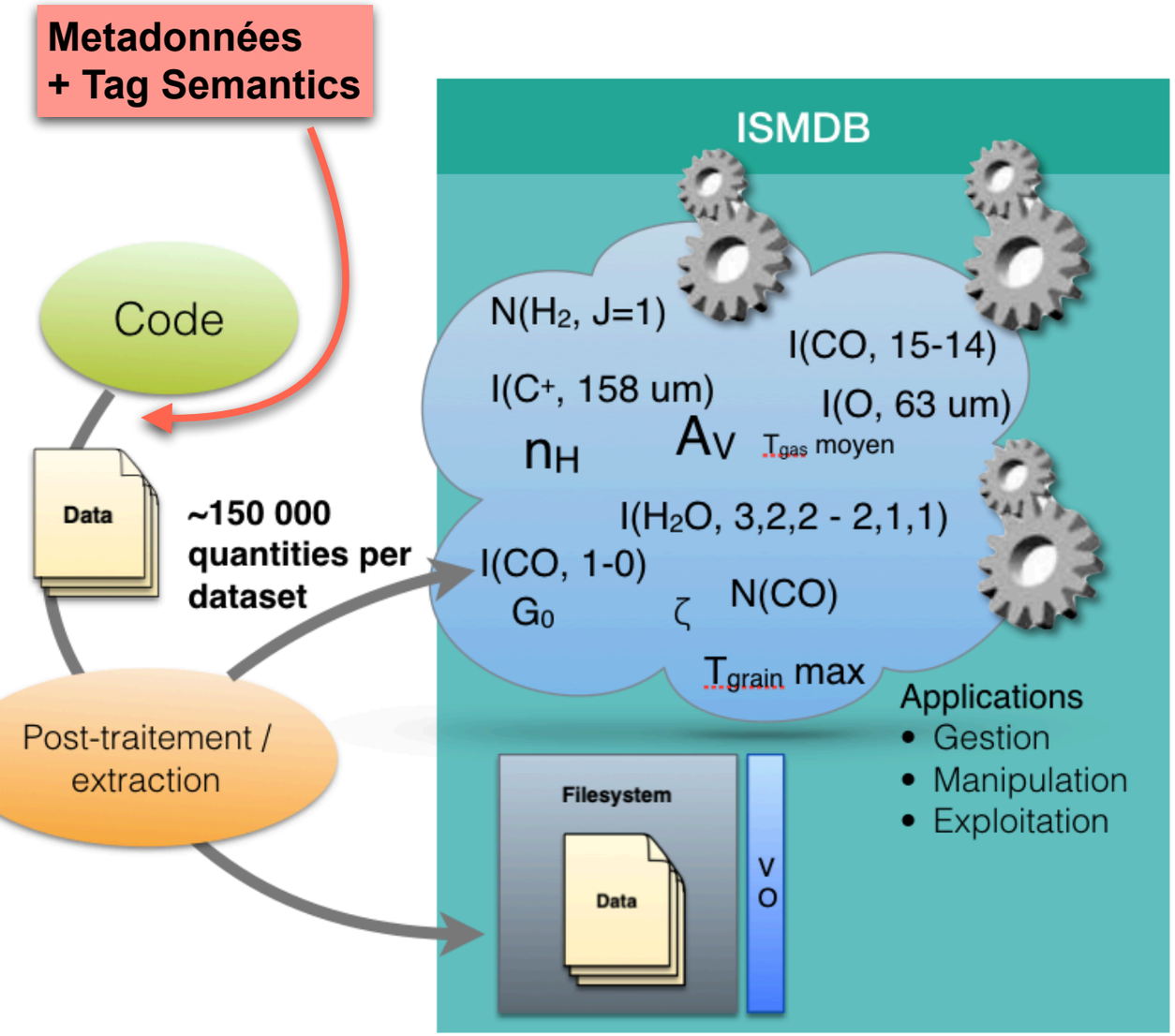


UV radiation field





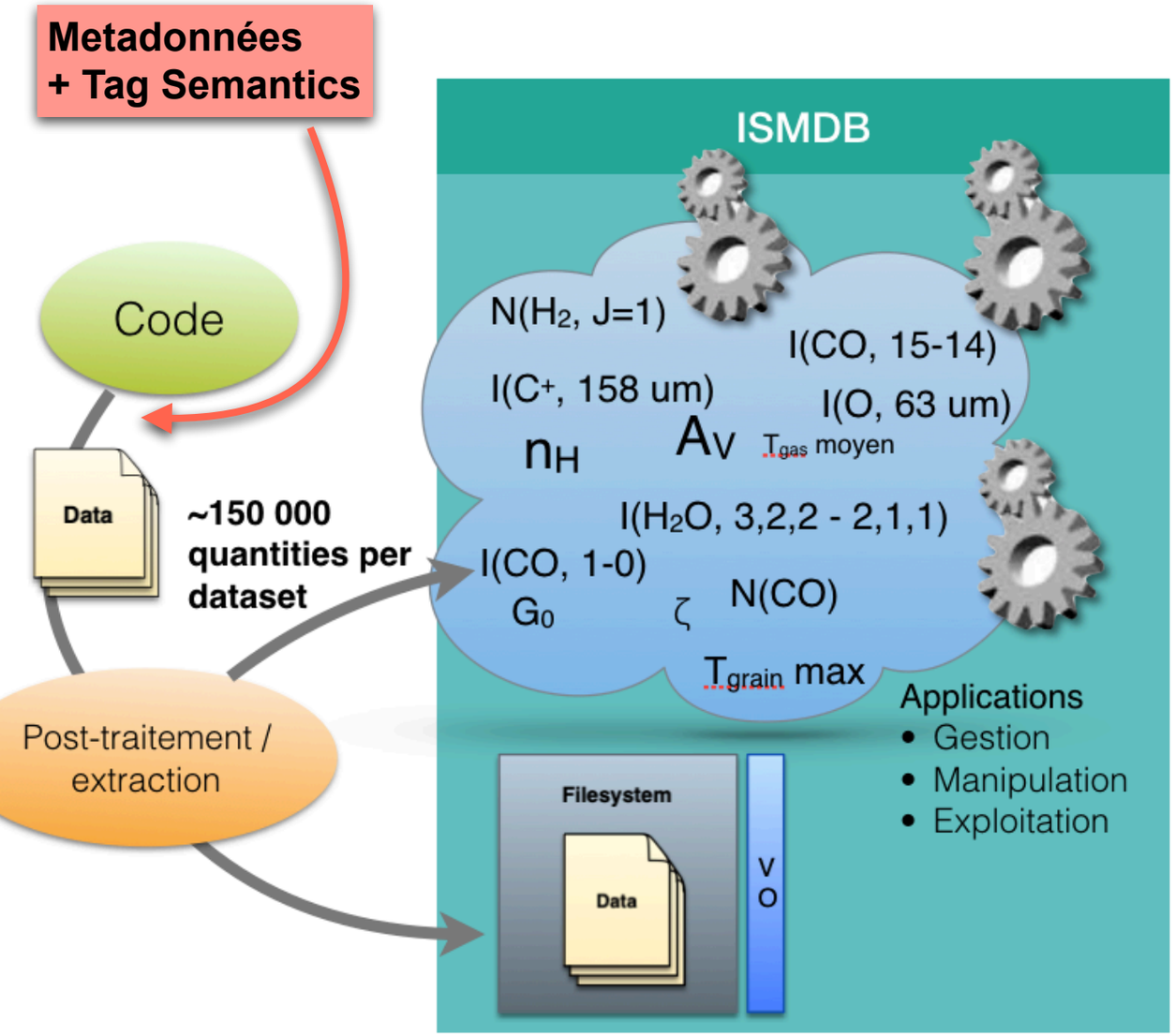
Application of vocabularies: ISMDB



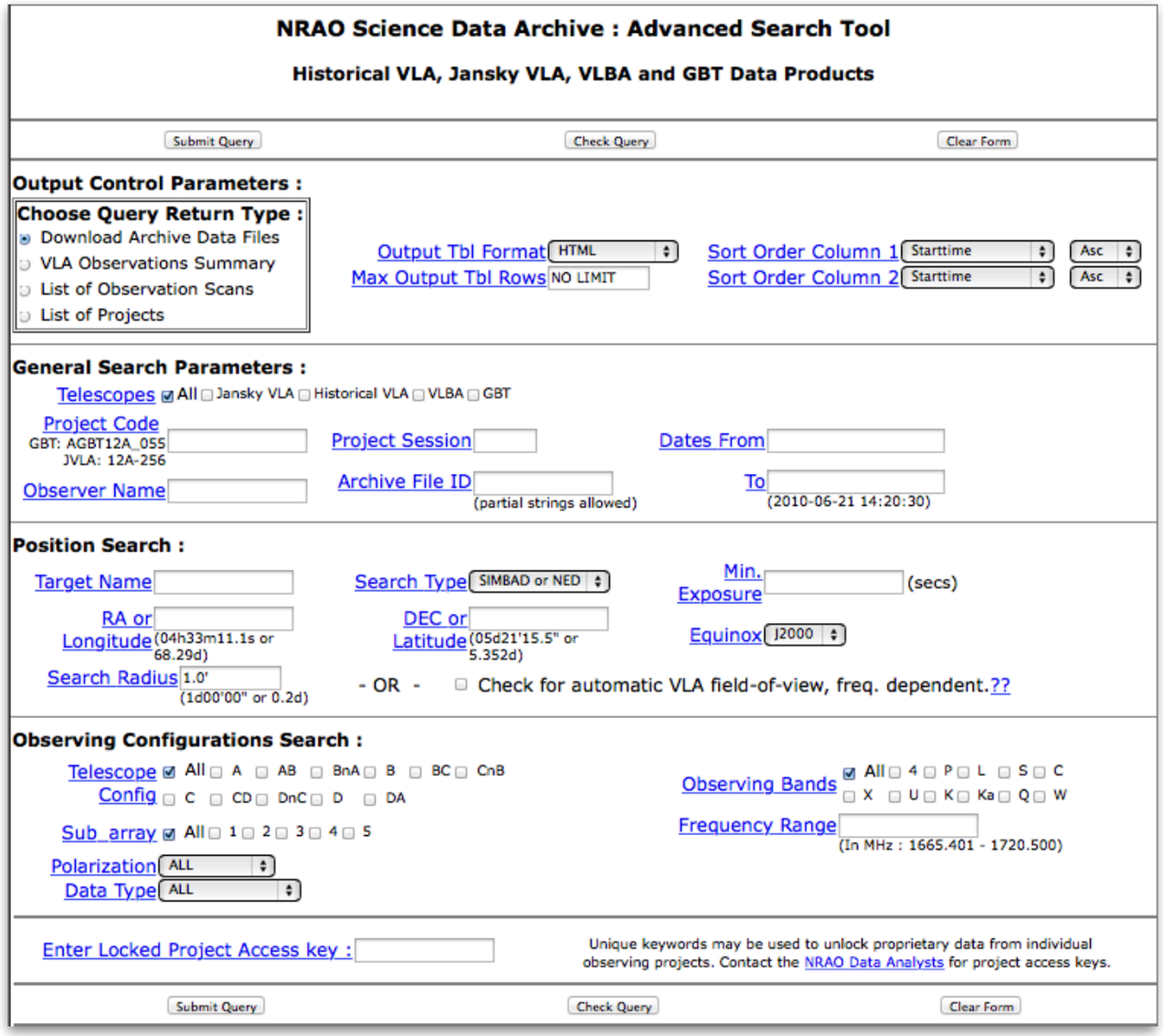
➔ **High dimension** database

➔ Interrogations possible on more than 150 000 quantities

Application of vocabularies: ISMDB



Example of a Classical interface (ex: VLA)

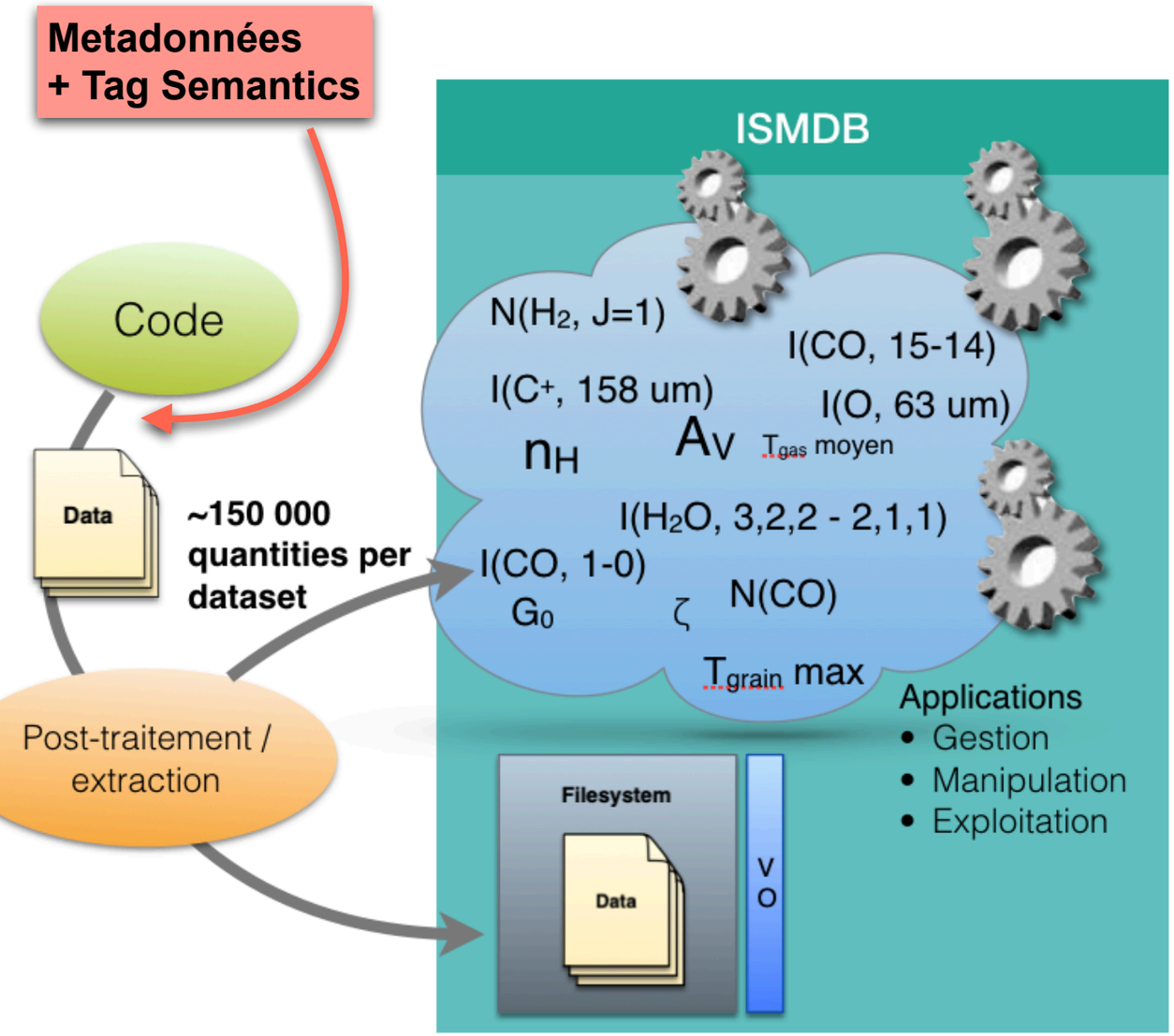


➔ **High dimension** database

➔ Interrogations possible on more than 150 000 quantities

For ISMDB, ~ 150 000 fields would be necessary !

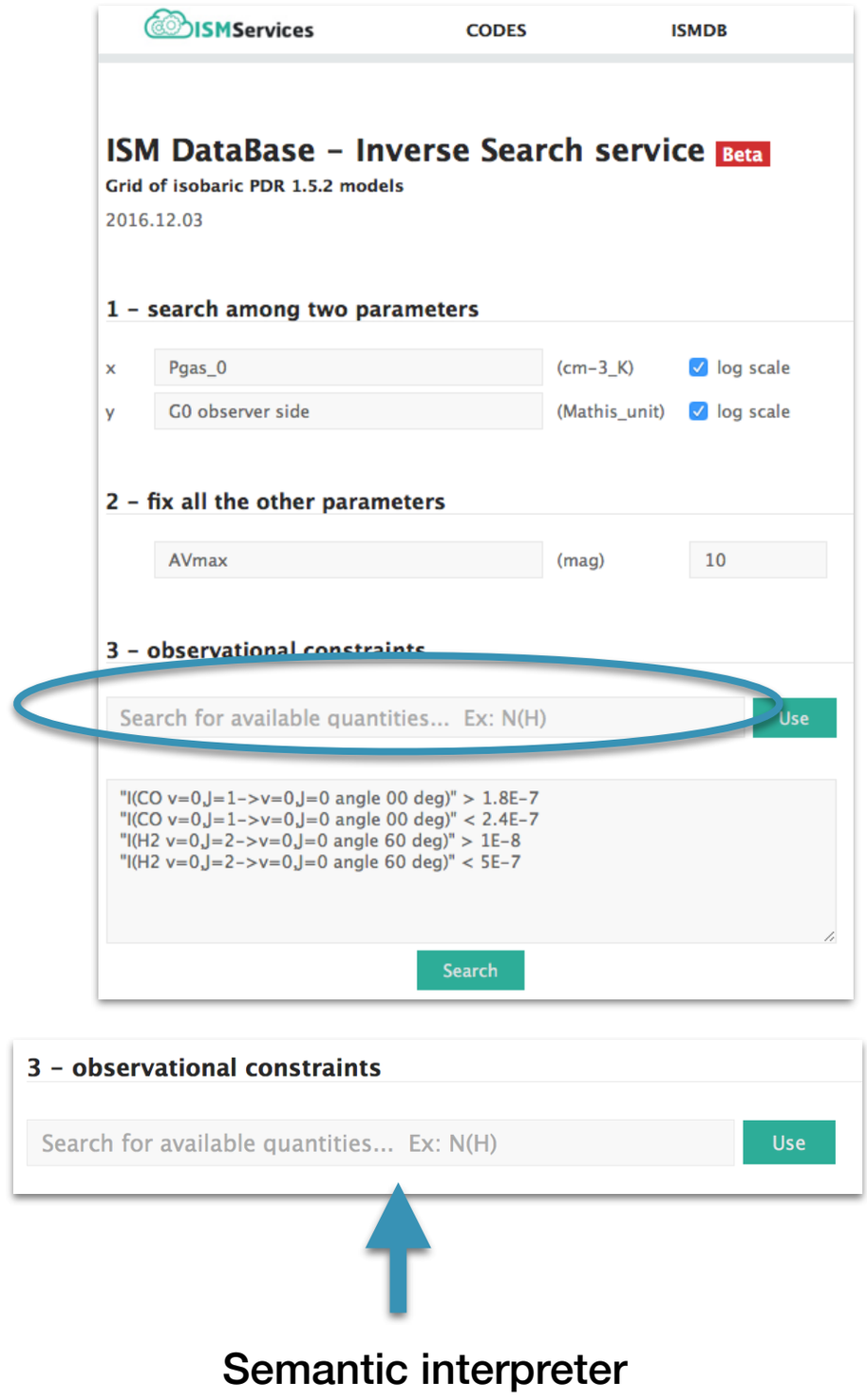
Application of vocabularies: ISMDB



➔ **High dimension** database

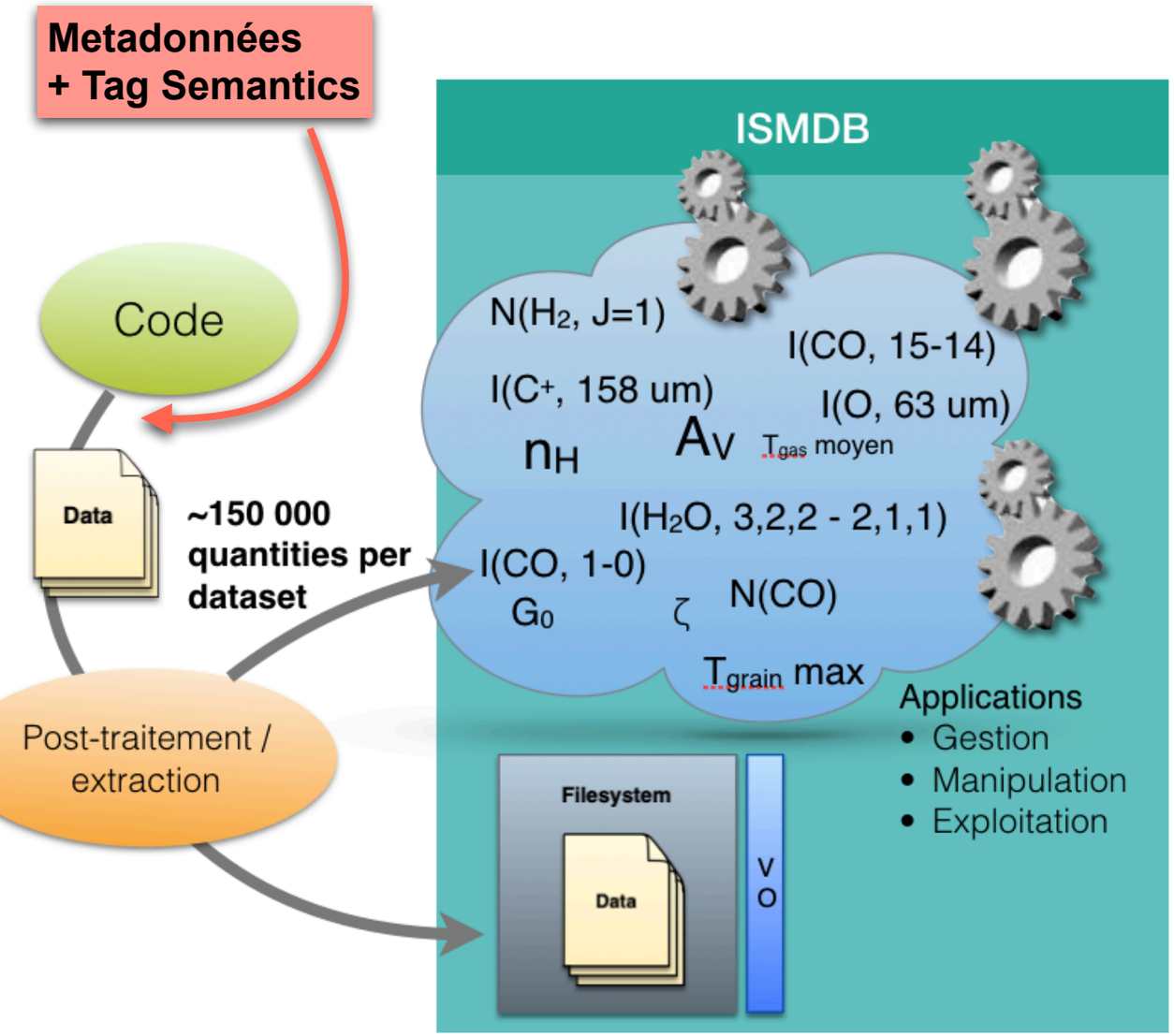
➔ Interrogations possible on more than 150 000 quantities

ISMDB interface

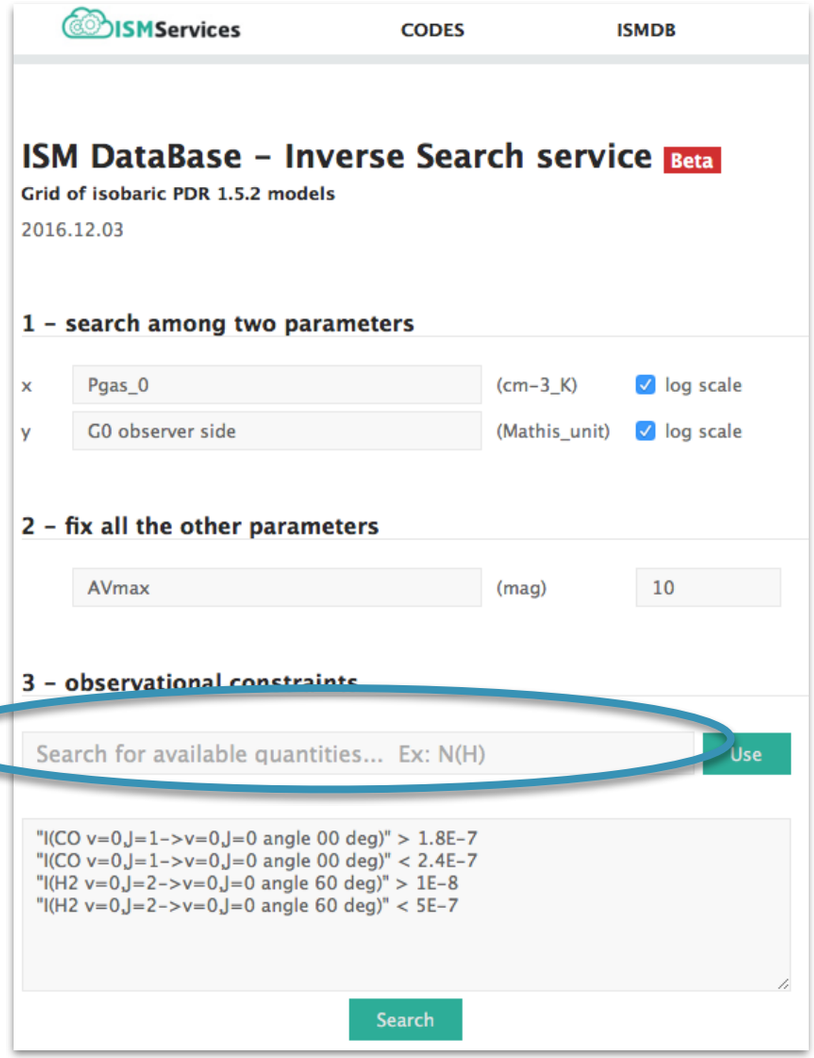


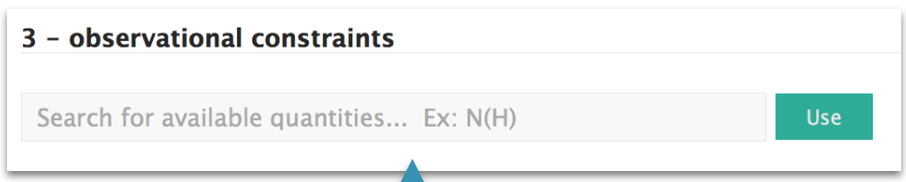
Semantic interpreter

Application of vocabularies: ISMDB



ISMDB interface





➔ **High dimension** database

➔ Interrogations possible on more than 150 000 quantities

tag each quantity with semantics concept

Semantic interpreter

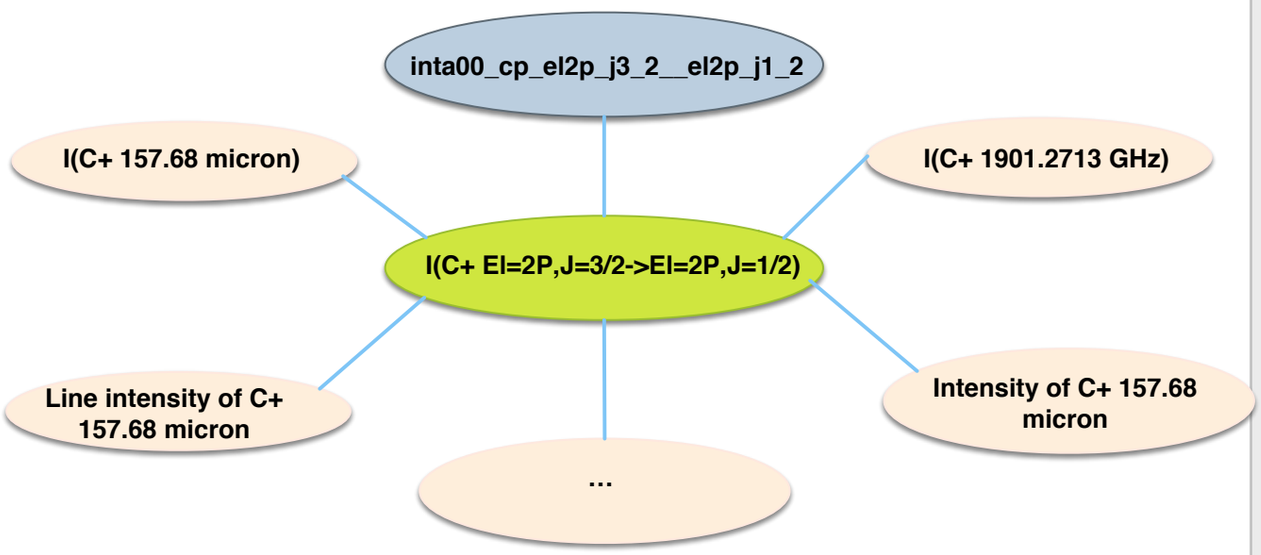
Application of vocabularies: ISMDB

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
- …

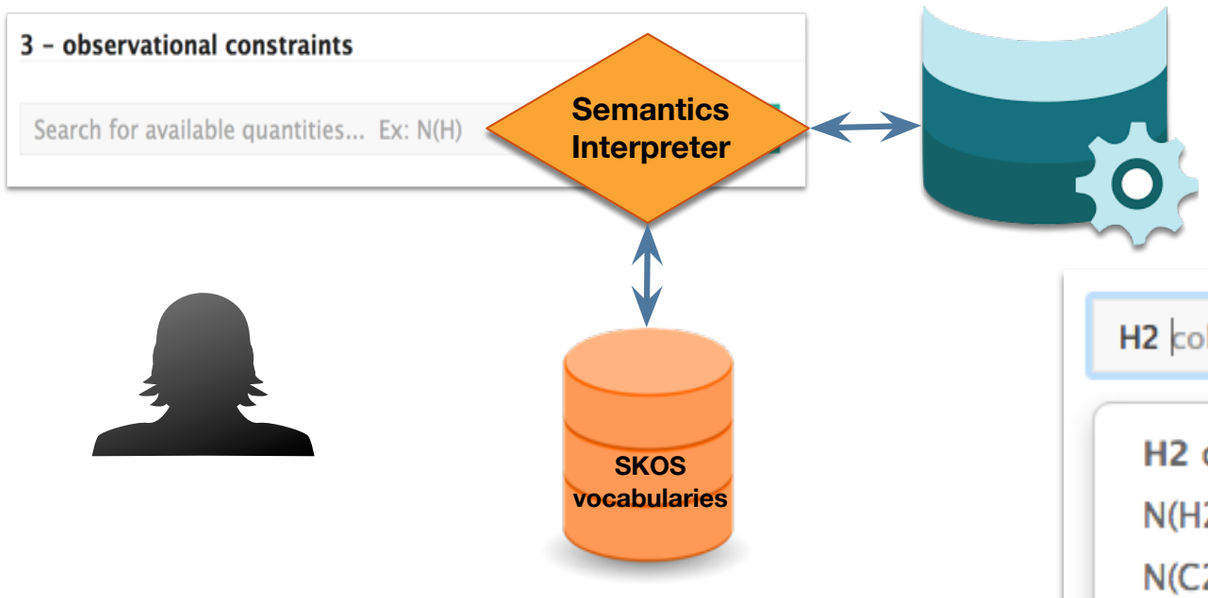
➔ specific vocabulary of ~ **300 000 concepts** with **synonyms**

<u>Example of the 157.7 micron C^+ line intensity</u>



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)