

1st SSHADE partners meeting

10-11 May 2016 – IPAG, Grenoble, France

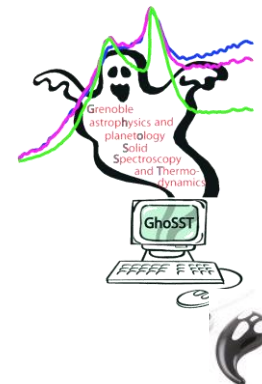
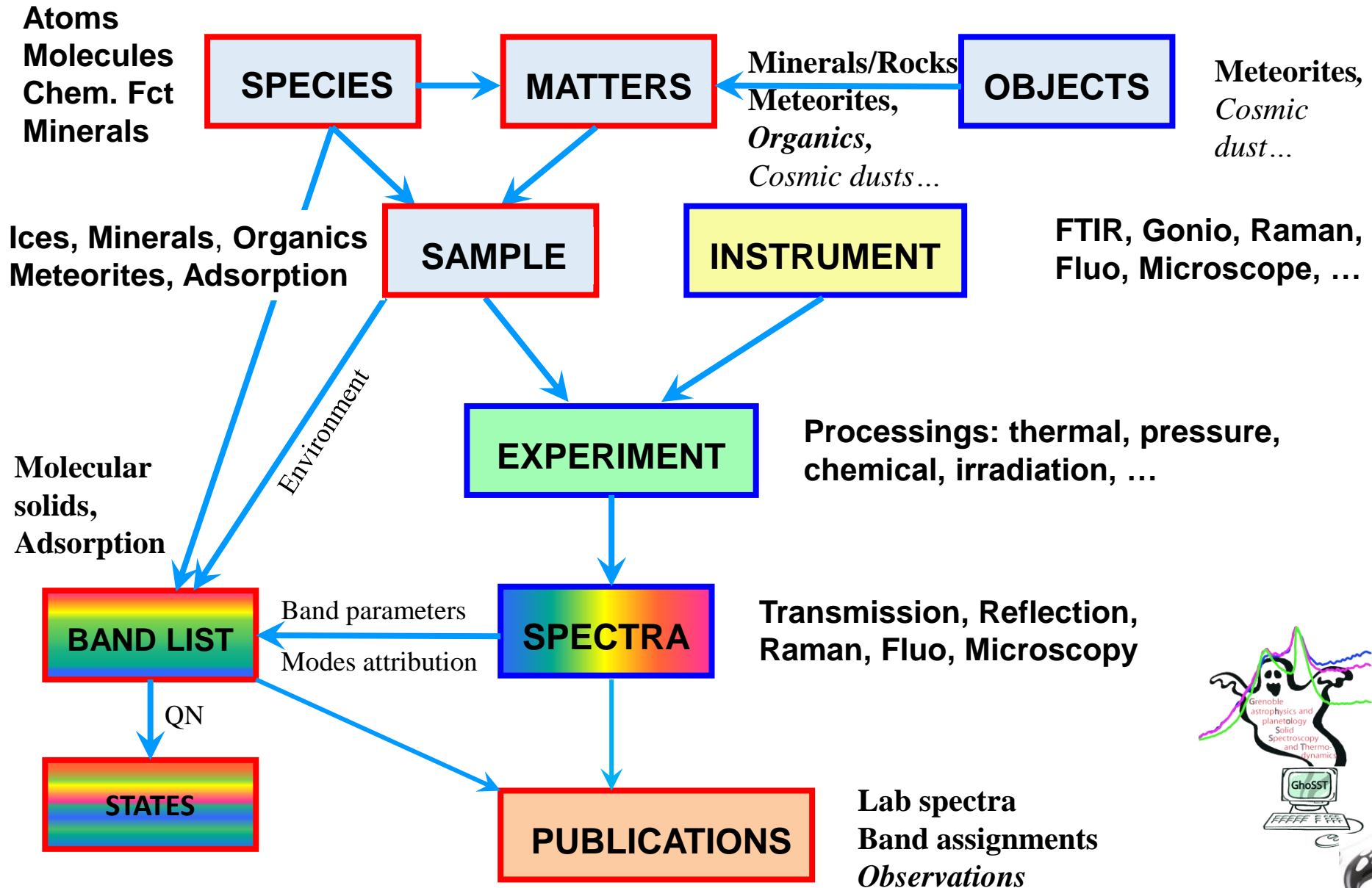
Manager training program

- Requisite tfor both scientific manager and database manager to get access to your SSHADE database
- 2 days of theoretical and practical training (in Grenoble or on-site)

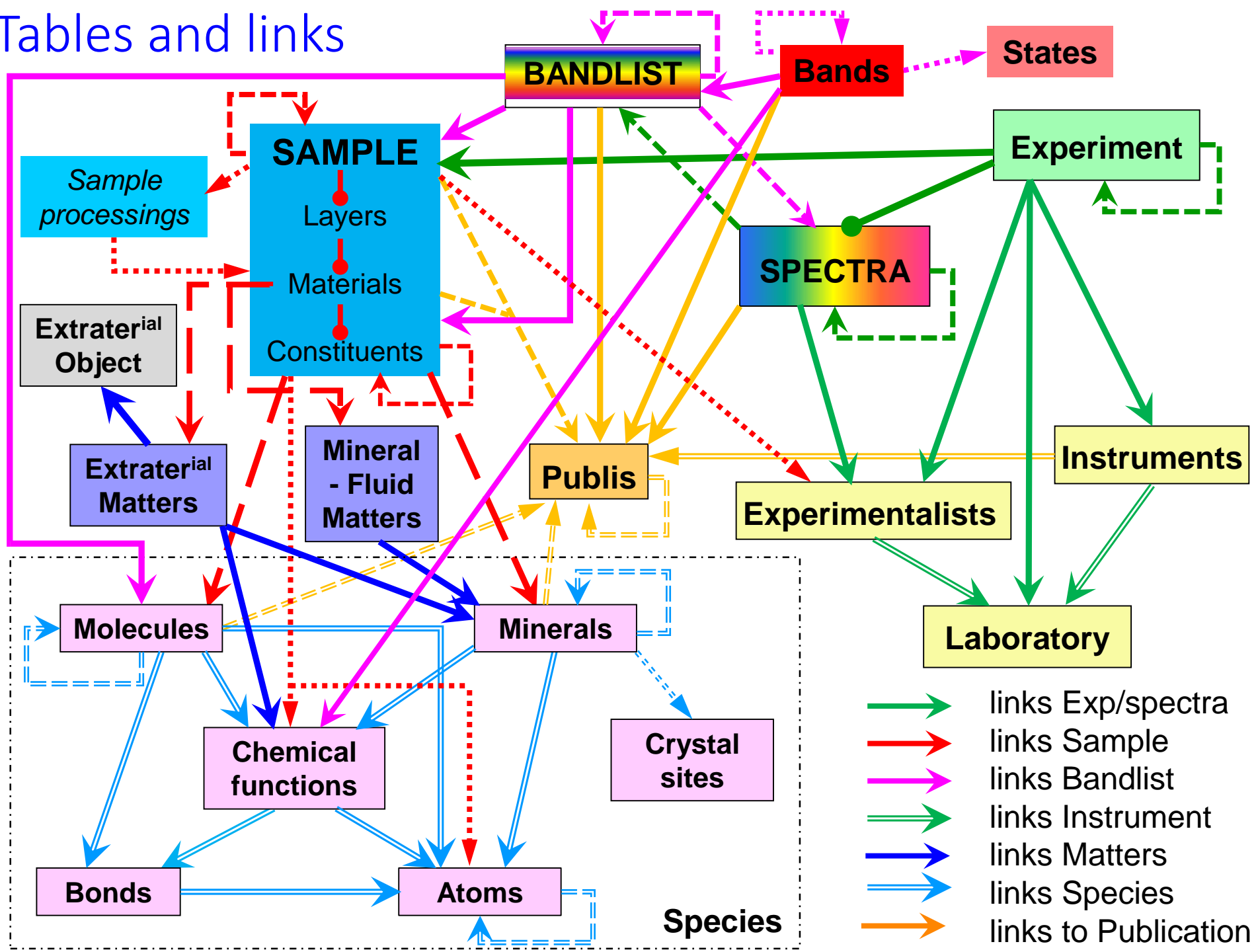
You will learn :

- The SSDM datamodel
- The different types of 'tables' and their links
- The different categories of Keywords
- Step by step process of data import
- The import xml files and how to fill them
- The import and validation tools. How to debug.
- How to manage and correct data
- The user interface

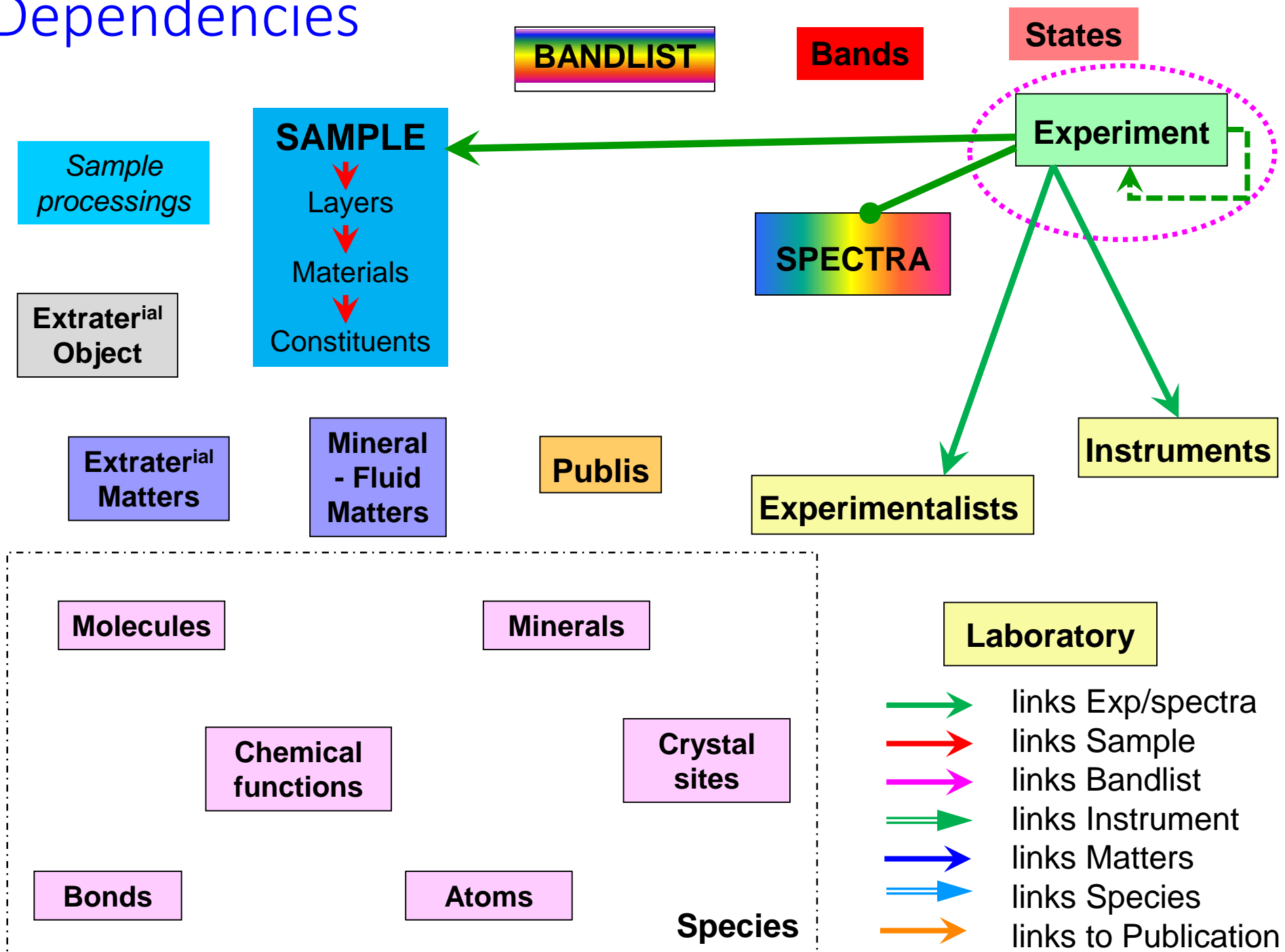
SSDM General Structure



Tables and links



Dependencies



Links: The « UID » concept

Experiment table

CREATE UID: experiment →

LINK UID: parent experiment →

LINK UID: copyright lab →

LINK UID: experimentalist(s) →

LINK UID: sample →

LINK UID: instrument →

LINK UID: instrument lab →

```
<?xml version="1.0" encoding="UTF-8"?>
<!--
Data type : Experiment and spectra
SSDM version: 0.5.5c
-->
<!-- Thin film CH4 crystalline II - deposited 15K - 0.545 µm -->
<import type="experiment_spectra">
  <experiment>
    <import_mode>first import</import_mode><!-- **M** {first import, ignore, no change, use existing}
    <uid>EXPERIMENT_BS_20130129_003</uid> <!-- **Mandatory to CREATE** Unique identifier code to the
    <!-- EXPERIMENT: HISTORY -->
    <date>1995-04-25</date> <!-- **M** 'YYYY-MM-DD' -->
    <parent_experiment_uid>EXPERIMENT_BS_20130129_002</parent_experiment_uid> <!-- Link to the UID
    <first_parent_experiment_uid>EXPERIMENT_BS_20130129_001</first_parent_experiment_uid> <!-- Link
    <!-- EXPERIMENT: COPYRIGHTS -->
    <copyright_laboratories>
      <laboratory><!-- multiple -->
        <uid>LABO_IPAG</uid> <!-- **M: or next KW** Link to the UID of the laboratory -->
      </laboratory>
    </copyright_laboratories>
    <experimentalists>
      <experimentalist_uid>EXPER_Frederic_Trotta_LGGE</experimentalist_uid><!-- multiple, Link to
      <experimentalist_uid>EXPER_Bernard_Schmitt_LGGE</experimentalist_uid>
    </experimentalists>
    <!-- EXPERIMENT: SAMPLE + INSTRUMENT -->
    <sample_uid>SAMPLE_BS_20130129_003</sample_uid> <!-- **Mandatory** Link to the UID of the samp:
    <instruments> <!-- **MANDATORY if no links with parent spectra** -->
      <instrument><!-- multiple -->
        <uid>INSTRU_Nicolet800_MIR</uid> <!-- **Mandatory** Link to the UID of the instrument--:
        <laboratory><!-- single -->
          <uid>LABO_LGGE</uid> <!-- **Mandatory** Link to the UID of the laboratory where the
          <acronym></acronym> <!-- **Mandatory: if no UID** -->
        </laboratory>
      </instrument>
  </experiment>
```

UNDERSTANDING YOUR SAMPLE STRUCTURE & COMPOSITION

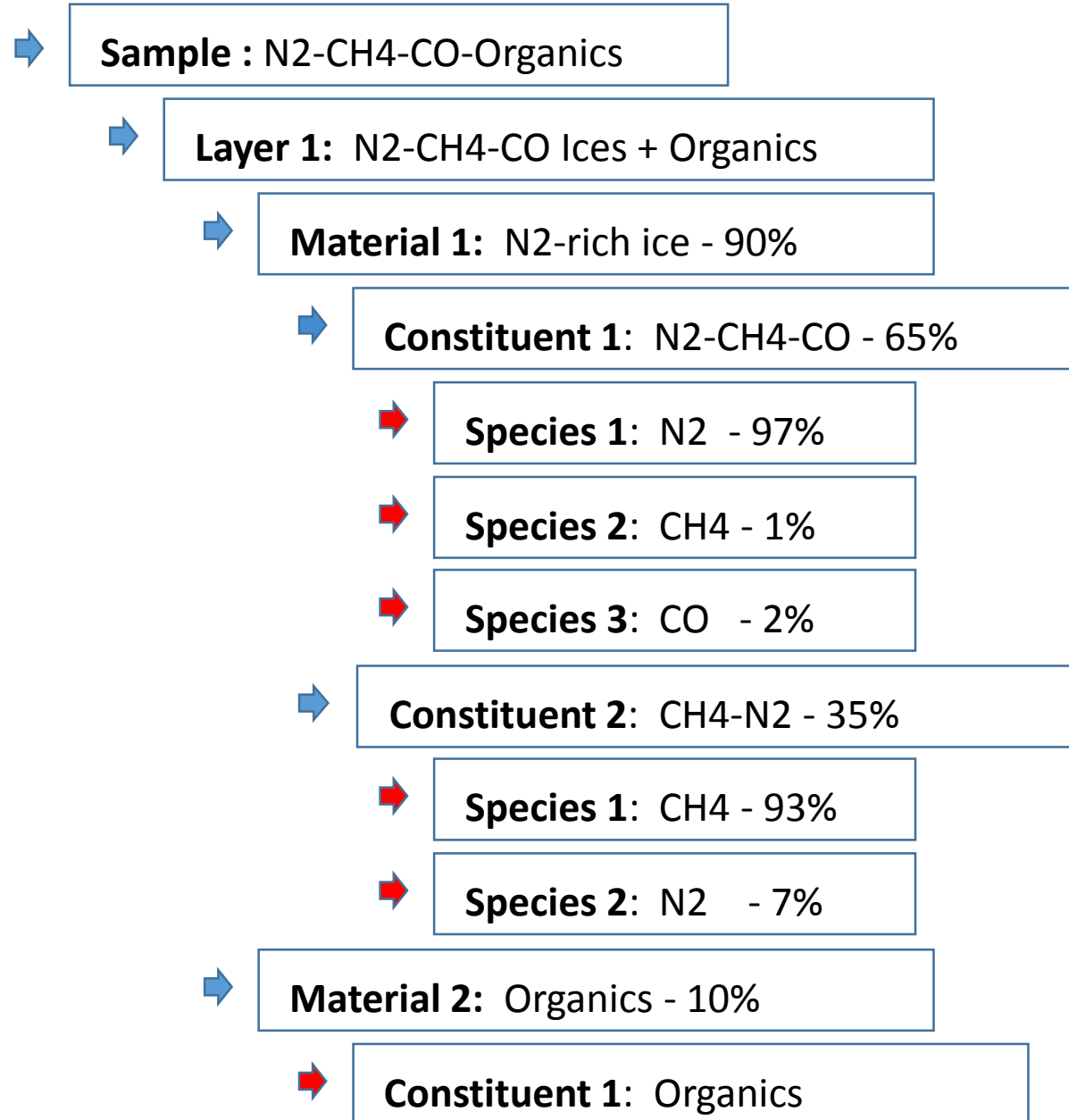
- **Which type of sample ?**

- Synthetic ? Using precursors ? Processings
- Natural ? Using a Matter ?

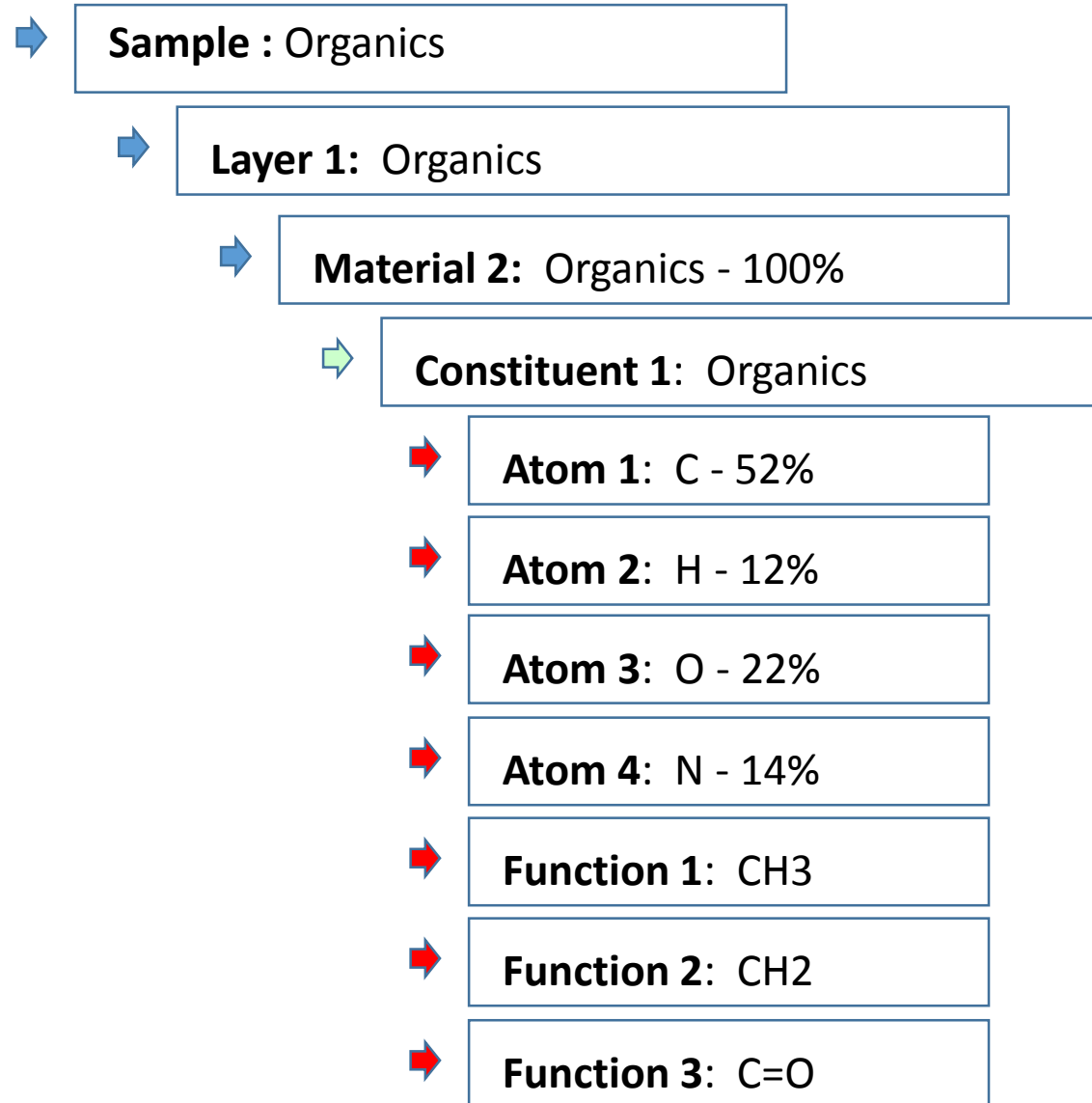
→ Describe 'on paper' the structure of your Sample

- Layers
- Materials (grains)
- Constituents (crystals)
- Species (molecules, minerals, or atoms, chemical functions)

Description of Sample structure



Description of Sample structure



PREPARING XML FILES FILLING

- **List all elements** you will need to link with your sample and experiment/spectra
 - **Verify their existence in the DataBase**
 - Laboratory
 - Instruments
 - Experimentalists
 - Species: Molecules, Minerals, (+ atoms, chemical functions)
 - Matters, Objects
 - Publications
- ➔ use « Producer / Full Search » in SSHADE
- ➔ take note of their UID

IMPORT TOOLS for DATA PRODUCERS

Developed:

- xml templates for each data type
- import tools + validators
- control interface
- import history
- import documentation
- import tutorials

GhosST
Back-end / Import / Sample

Import file

*Processed import file : sample_H2O_multilayer-demo_v6.xml.
*Import type : Sample.
***Database indexes** : (references are shown in *italico*)
 Sample : CO2+13C16O2+H2O / H2O amorphous - dep 10K - bi-film 7+2µm . Index : 58
 Layer n°1 : 1 materials.
 Material : H2O amorphous la. Index : 125
 Composition : 1 constituant.
 Precursor : H2O gas. Index : 126
 Composition : 1 constituant.
 Layer n°2 : 1 materials.
 Material : CO2+13C16O2+HD160 amorphous mix. Index : 127
 Composition : 1 constituant.
 Precursor : CO2+13C16O2+HD160 gas mix. Index : 128
 Composition : 1 constituant.
*Import finished.

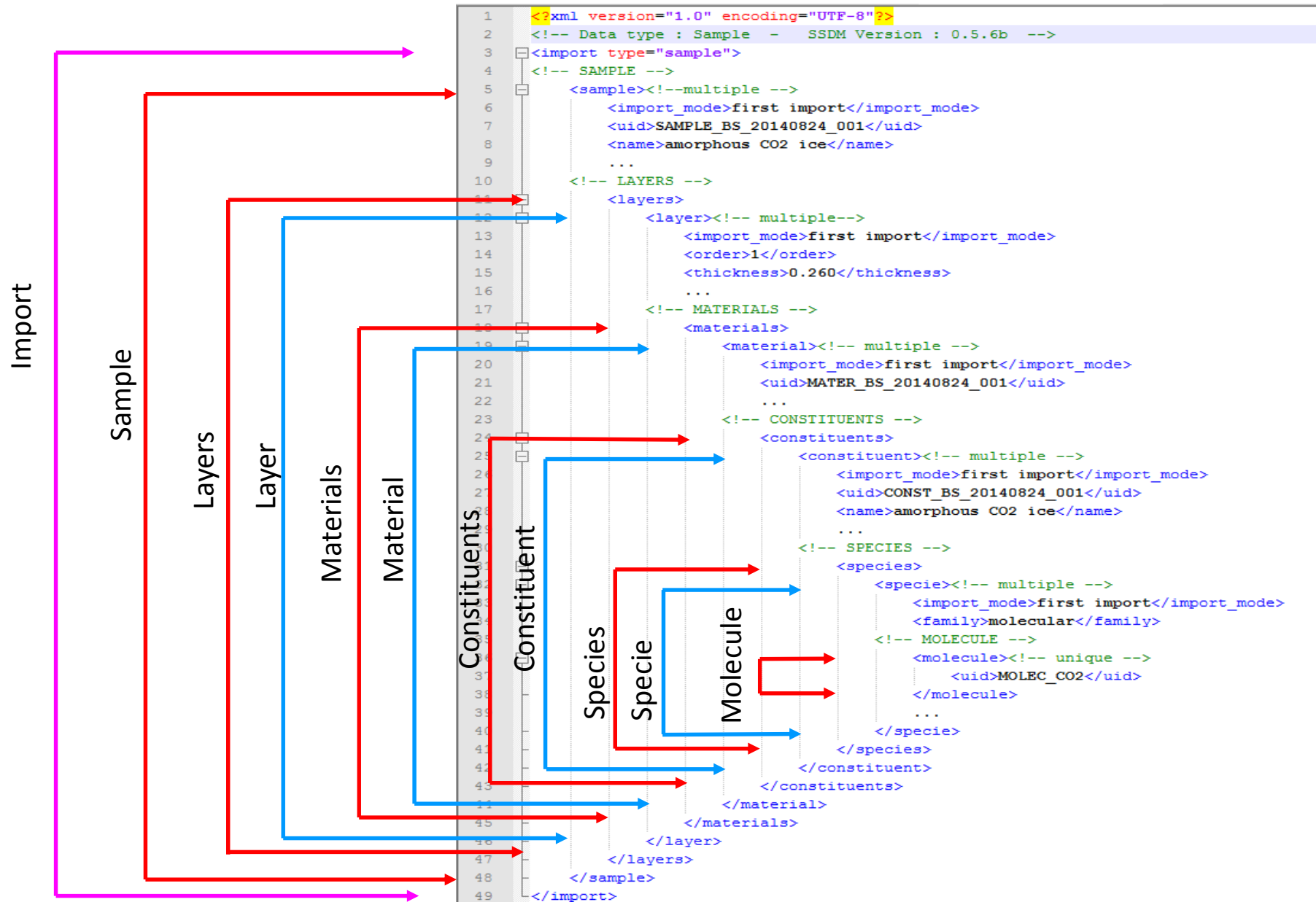
[Help](#)

PREPARING XML FILES + DATA IMPORT

Sample information

- **Describe the state and composition of your sample in xml file**
 - Type and links (parent sample, publication(s))
 - Layers / Materials / Constituents / Species
 - eventually use Matters
- **Option: You can give details about preparation protocol of your sample**
 - Precursor matters or species
 - Process(es) occurring a each steps
 - Change(s) occurring a each steps
- **Validate the xml**
 - **⚠ Simulation mode ⚠**
(until **Import finished** or message "Data file not joined 'file name' - Sample 'UID': Error during import of data")
- **possibly Zip with image files**
 - and continue to validate the xml in **⚠ Simulation mode ⚠** until you get **Import finished**
- **Import the sample**
 - verify all is OK using Producer / Full Search / Sample to search and visualize it
 - ✓ Note: your sample xml file will be also stored 'as it is' in the DB and retrievable with the 'import history' tab

XML: General structure



Templates: Mandatory KeyWords

- **Mandatory levels**

- Absolute mandatory: (needed to make working the database) **no Jocker**
Ex: `<uid>OBJMET_ALHA77307</uid> <!-- **Absolute Mandatory** UID: 'OBJMET_MeteoriteName' -->`
Ex: `<name>ALH A77307</name> <!-- **Absolute Mandatory** Name of the meteorite object -->`
- Mandatory: (needed for search or major info)
Ex: `<temperature></temperature> <!-- **Mandatory** Effective temperature of the sample during spectrum recording. [Float] Unit: in "sample_temperature_unit" --> Jocker = 'NULL'`
- Mandatory but with a default value (needed for search or major info)
Ex: `<import_mode>first import</import_mode> <!-- **Mandatory: Default=first import**-->`
- Optional mandatory (not needed for search in some cases)
Ex: `<hydration_number_range>2-6</hydration_number_range> <!-- **Mandatory for hydration series** Range of number 'n' of H2O in the mineral series -->`
- Free (not needed for search, simple info) **Jocker = ''**
Ex: `<secondary_name>vitriol</secondary_name> <!-- Alternative name used for the molecule -->`
- Calculated (not in the xml, calculated during import)
Ex: `<ID> </ID> <!-- *Calculated* Internal identifier -->`
Ex: `<sample_thickness> <!-- *Calculated from layer thicknesses* Sample thickness -->`

Templates: Typical Errors (2)

- **Bad data format:**

Wrong: `<temperature>100,0</temperature>` `<!-- sample temperature. (float) -->`

Good: `<temperature>100.0</temperature>`

Tip: copy/paste from the Enum/Open Enum attribut list

- Element 'temperature_max': '100,0' is **not a valid value of the union type 'float'**

- **Unexisting/error in UID link:**

Wrong: `<uid>PUBLI_Quirico_1895</uid>` `<!-- Link to the UID of the publication -->`

Good: `<uid>PUBLI_Quirico_1995b</uid>`

Tip: found and copy/paste UID with Menu Producer/Full search/Publications

- Unknown publication uid : PUBLI_Quirico_1895.**Unknown publication PUBLI_Quirico_1895.**

- **Missing Mandatory KW:**

Wrong: `<family></family>` `<!-- **Mandatory** Material family. Enum: { atomic, molecular, mineral, mixed molecular-mineral, mineral matter, fluid matter, organic-carbonaceous matter... -->`

Good: `<family>molecular</family>`

Tip: check ****Mandatory**** in comments

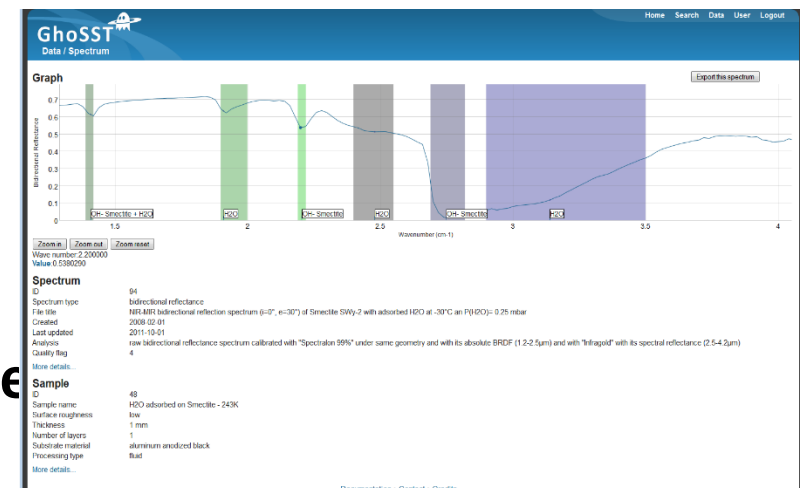
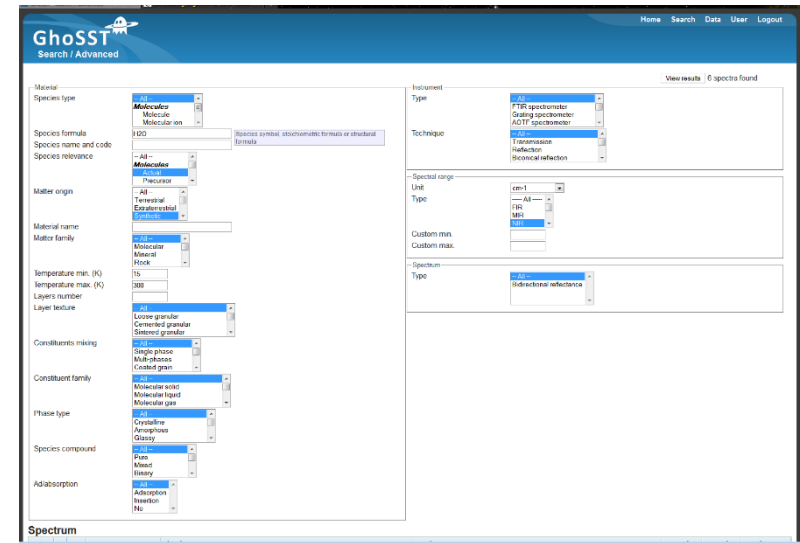
XML Template: filing data information

```
E:\BdD\templates XML\last XML\template_laboratories_commented_v054a.xml - Notepad++
Fichier Édition Recherche Affichage Encodage Langage Paramétrage Macro Exécution Compléments Documents ? X
bugs-GhoSST-janvier13tmp.txt reunions Corolys.txt experiment_spectra_test_v054c.xml template_laboratories_commented_v054a.xml
1 <?xml version="1.0" encoding="UTF-8"?>
2 <!--
3 Data type : Laboratory
4 SSDM version: 0.5.4a
5
6 Specific notes :
7 -
8
9 General notes :
10 - Most of the tags are optional, remove unnecessary ones.
11 - if the value (txt or numeric) of one tag is not known but wait a MANDATORY value,
12 then put 'NULL' and write a comment to keep track of the missing value. Remove
13 comment when value is added.
14 - Enumeration type must contain one item from the list given in brackets.
15 - Tags marked as multiple can be copied (with its block of sub-tag) if needed.
16 - UID (unique identifier) are references to other table in GhoSST, you have to found
17 this UID in the database beforehand (use Provider/Full Search).
18 - use a CDATA tag when a value contains one or more special characters (ie: &, >,
19 <,...). Example: <![CDATA[AT&T]]> for AT&T
20 -->
21 <import type="laboratory">
22 <laboratory><!-- multiple -->
23 <import_mode>first import</import_mode><!-- **Mandatory: Default=first import**
24 Mode of import of the 'laboratory' data. Enum: (first import, ignore, no change,
25 correction) -->
26 <uid>LABO_</uid><!-- **Mandatory to CREATE** Unique identifier code given to the
27 laboratory: 'LABO_LaboAcronym' -->
28 <acronym></acronym> <!-- **Mandatory** Acronym of the laboratory -->
29 <name></name> <!-- **Mandatory** Full name of the laboratory -->
30 <organization></organization> <!-- **Mandatory** Name of the organization to
31 which belong the laboratory -->
32 <address></address> <!-- **Mandatory** Address of the laboratory (without
33 country) -->
34 <country></country> <!-- **Mandatory** Country of the laboratory. XXX BS 046a
35 CHANGED Enum: {Australia, Austria, Belgium, Canada, China, Czech Republic,
36 Denmark, Finland, France, Germany, Greece, Hungary, India, Israel, Italy, Japan,
37 Republic of Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal,
38 Romania, Russian Federation, Spain, Sweden, Switzerland, Taiwan, Ukraine, United
39 Kingdom, United States, Venezuela, ...} -->
40 <urls><!-- Web address of the laboratory and organization -->
41 <url><![CDATA[]]></url><!-- multiple --> <!-- URL of web site -->
42 </urls>
43 <comments><![CDATA[]]></comment> <!-- Additional information on the laboratory
44 (Tel, ...) -->
45 </laboratory>
46 </import>
eXtensible length: 2250 lines: 31 Ln:19 Col:16 Sel:0|0 Dos\Windows ANSI as UTF-8 INS
```

```
E:\BdD\templates XML\last XML\template_laboratories_commented_v054a.xml - Notepad++
Fichier Édition Recherche Affichage Encodage Langage Paramétrage Macro Exécution Compléments Documents ? X
experiment_spectra_test_v054c.xml template_laboratories_commented_v054a.xml
1 <?xml version="1.0" encoding="UTF-8"?>
2 <!--
3 Data type : Laboratory
4 SSDM version: 0.5.4a
5
6 Specific notes :
7 -
8
9 General notes :
10 - Most of the tags are optional, remove unnecessary ones.
11 - if the value (txt or numeric) of one tag is not known but wait a MANDATORY value,
12 then put 'NULL' and write a comment to keep track of the missing value. Remove
13 comment when value is added.
14 - Enumeration type must contain one item from the list given in brackets.
15 - Tags marked as multiple can be copied (with its block of sub-tag) if needed.
16 - UID (unique identifier) are references to other table in GhoSST, you have to found
17 this UID in the database beforehand (use Provider/Full Search).
18 - use a CDATA tag when a value contains one or more special characters (ie: &, >,
19 <,...). Example: <![CDATA[AT&T]]> for AT&T
20 -->
21 <import type="laboratory">
22 <laboratory><!-- multiple -->
23 <import_mode>first import</import_mode><!-- **Mandatory: Default=first import**
24 Mode of import of the 'laboratory' data. Enum: (first import, ignore, no change,
25 correction) -->
26 <uid>LABO_IPAG</uid><!-- **Mandatory to CREATE** Unique identifier code given to
27 the laboratory: 'LABO_LaboAcronym' -->
28 <acronym>IPAG</acronym> <!-- **Mandatory** Acronym of the laboratory -->
29 <name>Institut de Planétologie et Astrophysique de Grenoble</name> <!--
30 **Mandatory** Full name of the laboratory -->
31 <organization>University J. Fourier, CNRS/OSUG</organization> <!-- **Mandatory**
32 Name of the organization to which belong the laboratory -->
33 <address>124 rue de la Piscine, Bâtiment D de Physique, 38042 Grenoble Cedex
34 </address> <!-- **Mandatory** Address of the laboratory (without country) -->
35 <country>France</country> <!-- **Mandatory** Country of the laboratory. Enum:
36 {Australia, Austria, Belgium, Canada, China, Czech Republic, Denmark, Finland,
37 France, Germany, Greece, Hungary, India, Israel, Italy, Japan, Republic of Korea,
38 Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian
39 Federation, Spain, Sweden, Switzerland, Taiwan, Ukraine, United Kingdom, United
40 States, Venezuela, ...} -->
41 <urls><!-- Web address of the laboratory and organization -->
42 <url><![CDATA[http://ipag.osug.fr]]></url><!-- multiple --> <!-- URL of web
43 site -->
44 </urls>
45 <comments><![CDATA[postal address: BP 53, F-38041 GRENOBLE Cedex 9, France.]]>
46 </comment> <!-- Additional information on the laboratory (Tel, ...) -->
47 </laboratory>
48 </import>
eXtensible length: 2474 lines: 31 Ln:29 Col:87 Sel:0|0 Dos\Windows ANSI as UTF-8 INS
```

How to verify your data in SSHADE

- **Go to**
 - **The user search interface**
 - or
 - **The producer 'full search' interface**
- Search for your data
- Go to the details pages
 - sample,
 - spectra, ...
- Check that:
 - all information is correct
 - the spectrum/a is/are displayed correctly
- **Take note of possible corrections**
 - and set <import mode> to 'correction'
- **Prepare your corrections in the xml file**
- **Re-import the Data in 'correction mode'**



How to correct xml files

- **Take your last xml file or Get it from SSHADE**
 - use Producer/Full Search/'type of data' (Ex: 'sample')
 - go to the page of the data and to the 'import history' tab
 - 'Get the file' of the last correction
- **Make your corrections**
 - and set <import mode> to 'correction'
- **Validate the xml corrections**
 - Browse to your xml file
 - verify that box 'Simulation mode' is checked
 - **⚠ Simulation mode ⚠**
 - correct the bugs until you get: **Import finished.**
- **Import the Data**
 - uncheck box 'Simulation mode'
 - zip with data file(s) only necessary if you want to change a file (new version)

Home Search Data Producer Manager User GhosST

Data / Molecule / Sample / Matter / Sample / Sample / Sample

Sample Processings Parameters Experiments Publications **Import history**

Import history

Show 10 entries Search:

Import ID	ID	Date	Producer name	Mode import	Import file
991	7636	2013-12-06	Bernard Schmitt	first import	Get the file(s)
1025	7761	2013-12-12	Bernard Schmitt	correction	Get the file(s)
1054	7918	2013-12-14	Bernard Schmitt	correction	Get the file(s)
1079	8136	2013-12-14	Bernard Schmitt	correction	Get the file(s)
1735	12772	2014-08-11	Bernard Schmitt	correction	Get the file(s)

Showing 1 to 5 of 5 entries

Documentation Contact History Credits Statistics

Home Search Data Producer Manager User GhosST

Producer / Import / Sample

Import file

Simulation mode

⚠ Simulation mode ⚠

*Processed import file : sample_HC3N_NIR_100K_v056a.xml.
*Import type : Sample.
*Database indexes :
Sample : SAMPLE_BS_20130206_001, Index : 307, Mode : correction.
Layer : LAYER_1_BS_20130206_001, Index : 282, Mode : correction.
Material : MATER_BS_20130206_001, Index : 1015, Mode : correction.
Constituent : CONST_BS_20130206_001, Index : 477, Mode : correction.

Import finished.

Data Copyright(s) and heritage

- **Copyright** is defined at the experiment level
=> laboratory at the time of data acquisition
- **Copyright** may be granted to multiple labs
(experimentalists labs, instrument lab, ...)
- **if a researcher move to another lab**
Just need to create another 'experimentalist' at the new lab with link(s) to previous lab he/she want to grant copyrights
=> **copyrights** from his/her experiments in the previous labs will be also attributed to his/her current lab

(note: we are still reworking these rules and links)

```
<!-- SCHMITT BERNARD - IPAG -->
<experimentalist><!-- multiple -->
  <import_mode>no change</import_mode>
  <uid>EXPER_Bernard_Schmitt_IPAG</uid>
  <!-- EXPERIMENTALIST NAME -->
  <firstname>Bernard</firstname>
  <lastname>Schmitt</lastname>
  <acronym>BS</acronym>
  <status>researcher</status>
  <years>2011-now</years>
  <!-- EXPERIMENTALIST LABORATORIES -->
  <laboratory>
    <uid>LABO_IPAG</uid>
    <acronym></acronym>
  </laboratory>
  <parent_laboratories>
    <laboratory><!-- multiple -->
      <uid>LABO_Leiden_obs</uid>
    </laboratory>
    <laboratory>
      <uid>LABO_LGGE</uid>
    </laboratory>
    <laboratory>
      <uid>LABO_LPG</uid>
    </laboratory>
  </parent_laboratories>
  <copyright_laboratories>
    <laboratory><!-- multiple -->
      <uid>LABO_IPAG</uid>
    </laboratory>
    <laboratory>
      <uid>LABO_LPG</uid>
    </laboratory>
  </copyright_laboratories>
  <!-- EXPERIMENTALIST CONTACTS -->
  <email>Bernard.Schmitt@obs.ujf-grenoble.fr</email>
  <phone>+33 (0)476 51 41 50</phone>
  <urls>
    <url><![CDATA[http://ipag.osug.fr/~schmitt]]></url>
  </urls>
  <comments><![CDATA[]]></comments>
</experimentalist>
```


Training of SSHADE partners (as May 2016)

- **IPAG / Planéto**, Grenoble - F (Bernard Schmitt, Lydie Bonal, Damien Albert)
- **Open University**, Milton Keynes – UK (Nigel Mason)
- **IAS**, Univ. Paris-Sud - F (Emmanuel Dartois, Donia Baklouti)
- **IRAP / GPPS**, Toulouse - F (Patrick Pinet, Yves Daydou)
- **IRAP / MICMAC**, Toulouse - F (Karine Demyk, Yves Daydou)
- **LPG**, Univ. Nantes - F (Yann Morizet, Manuel Giraud)
- **Space and Planetary Science Division**, Univ. of Bern - CH (Antoine Pommerol)
- **PIIM**, Univ. Aix-Marseille - F (Patrice Theulé)
- **IAPS**, INAF, Roma - I (Alessandra Rotundi, Vincenzo della Corte)
- **IAPS**, INAF, Roma - I (Fabrizio Capaccioni, Christian Carli)
- **LISA**, Univ. Paris-Est - F (Nicolas Fray)
- **AIU Observatory**, Jena - D (Harald Mutschke, Jürgen Weiprecht) [DOCCD 'database']
- **Centro de Astrobiología**, INTA-CSIC – E (Guillermo Muñoz Caro)
- **Instituto de Estructura de la Materia**, Madrid – E (Vicente Timón, Miguel Angel Moreno)
- **LATMOS / IMPEC**, Institut Pierre Simon Laplace - F (Nathalie Carrasco)
- **LGL / ENS-Lyon** - F (Bruno Reynard, Gilles Montagnac (exp.), Razvan Caracas (theory))
- **Konkoly Astronomical Institute** – HU (Akos Kereszturi)
- **Planetary Geology Lab.**, Institute of Geological Sciences, Polish Academy of Sciences – PL (Joanna Gurgurewicz, Luigi Castaldo)
- **Clay Mineral Laboratory**, Institute of Geological Sciences, Polish Academy of Sciences – PL (Joanna Gurgurewicz, Luigi Castaldo)
- **ESRF / FAME line**, Grenoble – EU / F (Denis Testemale, Isabelle Kieffer)



Planning of database managers training (to be discussed)

- 2014 IAS+PIIM+LPGN (2 SM + 2 DBM)
- 2014 IRAP GPPS+MICMAC (F) (2 SM + 2 DBM)
- Dec. 2016 IEM (E) (1 SM)
- May 2016 AIU(G) + CML + PGL (PL) (2 SM + 2 DBM)
- Oct/Nov. 2016 ESRF (F) + Bern (CH)
- Dec. 2016 'recyclage' 5 french groups
- Feb. 2017 IAS + LPGN + IEM (complement)
- April 2017 IAPS (I: 2 groups)
- October 2017 OU (GB) + CAB (E)
- April 2018 AGC (HU) + LGL (F)
- October 2018 LISA (F) + LATMOS (F)

SSHADe Time line

2015

- September
- November
- December

Start of Europlanet 2020-RI

Start of SSHADe development

Training database managers (session #0)

2016

- May
- May
- August
- November

1st SSHADe meeting

Training database managers (session #1) **(D6.3 VAA VESPA – Y1)**

SSHADe prototype delivery **(D11.5 JRA VESPA)**

Training database managers (session #2)

2017

- February
- April -May
- June
- August
- September
- October
- October
- October

Training database managers (session #3)

Training database managers (session #4)

3 databases ingested in SSHADe **(D6.3 VAA VESPA – Y2)**

SSHADe infrastructure **(D11.7 JRA VESPA)**

Training users EPSC (session #1) **(D6.5 VAA VESPA)**

6 databases ingested in SSHADe

2nd SSHADe meeting

Training database managers (session #5)

SSHADE Time line

2018

- March Training database managers (session #6)
- April Training users EGU (session #2) (D6.5 VAA VESPA)
- June 9 databases ingested in SSHADE (D6.3 VAA VESPA – Y3)
- September Training users EPSC (session #3) (D6.5 VAA VESPA)
- October Training database managers (session #7)
- November 12 databases ingested in SSHADE

2019

- March 15 databases ingested in SSHADE
- March 3rd SSHADE meeting
- April Training users EGU (session #4) (D6.5 VAA VESPA)
- June 18 databases ingested in SSHADE
- August SSHADE with 18-20 databases (D6.3 VAA VESPA – Y4)
- August End of Europlanet 2020-RI