SPICA-VIS

Observation Database Requirements

Authors: Nicolas Nardetto, Denis Mourard, Karine Perraut

Doc. No. : SPICA-VIS-0021

Issue: 1.4

Date: 19/10/2020

CHANGE RECORD

ISSUE	DATE	SECTION	COMMENTS
1.0	10/09/2020	All	Creation DM/NN
1.1	20/09/2020	Sect. 3	DM: more details after JMMC feedbacks – Program/Night/Run
1.2	21/09/2020	Sect. 3	DM+KP: description of the tools; clarification of the night scheduling scheme.
1.3	25/09/2020	Sect. 3 & 4	DM+NN. Addition of section 4 and update of section 3 on the night scheduling
1.4	19/10/2020	Sect. 3	DM

Table of Contents

1. Scope

The scope of this document is to describe the database considered for the management of the observations with CHARA/SPICA and for the scientific exploitation of the large survey program. This database is called SPICA-DB.

2. Context and relevant documents

One of the main goals of the CHARA/SPICA instrument (Mourard et al. 2018) is to observe a sample of ~1000 stars over three years and to produce a public catalogue of stellar and planetary parameters. The description of the corresponding scientific objectives has been done in the Science Survey Management (SSM) Document (SPICA-VIS-0020). Many other observing programs will also be realized on the time open for the CHARA collaboration and for the general community. With a perspective of more than 100 nights per year, multi-semester programs, and the critical need for a control of the homogeneity of the final public catalogue, a careful management of the observations, their preparation, their execution, and their analysis is therefore mandatory. The interface with the SPICA Observing Software has been described in document SPICA-VIS-003. For the survey considerations, the SSM document has identified the different actions that will be needed up to the creation of the public catalogue.

We propose the development of a database called SPICA-DB. It contains the list of targets, of calibrators, and the observing blocks. SPICA-DB will contain a few thousands of entries. A representation of the database is proposed in Figure 1. We distinguish three steps: 1/ Preparation with, 2/Operation, and 3/ Exploitation. The preparation step corresponds to the operations needed for feeding the database with the list of targets, calibrators (either from the survey of from the open-time programs). It is based on the SPICA-DFS Software (SPICA Database Feeding Software). In this step, consultations are also possible. In the step 2 (operation), the database is used for the night scheduling process (SPICA-NSS), for collecting the results of the automatic data reduction software (SPICA-DRS), and the results of the Quality Control software (SPICA-QCS). Finally, the step exploitation uses the database for the science analysis and for the extraction of the final science products (SPICA-SAS + extraction of catalogues). These pieces of software are interfaced to the Database by various dedicated API.

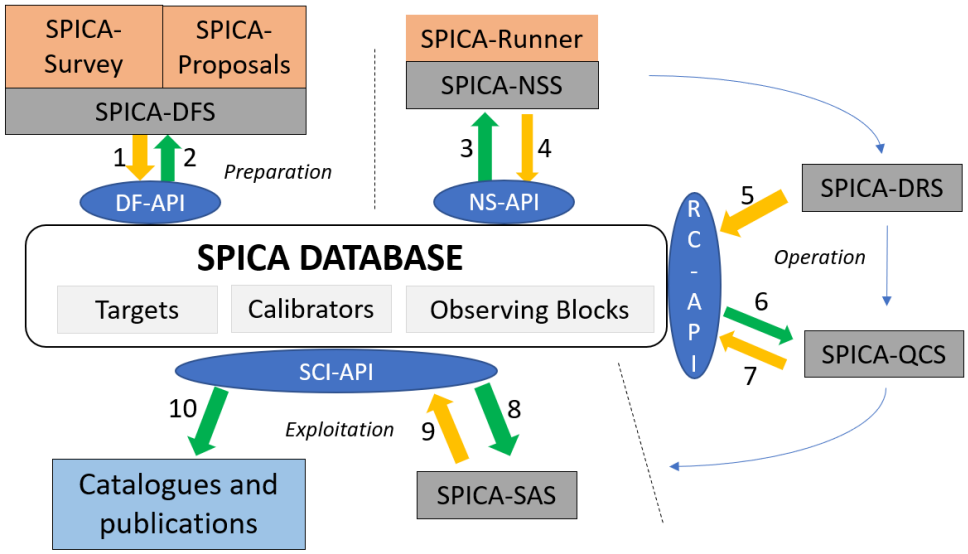

Figure 1: Schematic representation of the database. The yellow (resp. green) arrows correspond to feeding (resp. querying) actions. The grey boxes represent different user's applications linked to the DB. The blue box represents the final product in term of the public catalogue for the survey program

3. Database Requirements

3.1 General characteristics

- University of Aarhus is proposing its help to duplicate the SONG database (https://soda.phys.au.dk/) and to adapt it for the SPICA purpose. Another solution is based on new developments on the basis of the OBSPORTAL base developed by the JMMC group in Grenoble.
- SPICA-DB will be hosted on a JMMC server with backup and mirror services. It should be accessible all the time, through an authentication system to be defined.
- The encoding of the database, as well as all the related tools, should be adapted to the creation of additional columns.
- Four levels of users are defined:
    - Admin: with all the rights associated to the management of the database. This account will permit the creation and configuration of new accesses to the database. It will also manage all the services associated to the database.
    - Manager: high-level users having access to all the information contained in the database. Read and Write permissions are granted.
    - User: this account is dedicated to persons being responsible of some science programs generating SPICA observations. Specific Read and Write permissions are granted.
    - Runner: this account is granted to persons responsible of observing runs. This is a read-only account.
- The database contains three tables: the list of targets, the list of calibrators, and the list of observing blocks. A target entry contains all the relevant information for its management in terms of scheduling, observation, and scientific exploitation.
- A target entry contains the following fields, organized in three levels:
    - Observing Parameters
        - Target ID (s)
        - RA, DEC
        - Magnitude V, R, J, H, K
        - Estimated angular diameter (UD_R, UD_H)
        - PI Name and Program Name
        - A series of fields containing the required information for the scheduling of the observation (priority, setup, calibrator's specifications, quality criteria…).
        - Information on the completion rate.
    - Astrophysical input parameters
        - A series of fields containing the various astrophysical information of the star, needed for the completion of the specific science goal. Details are given in Section 4.
    - Astrophysical output parameters
        - A series of field containing the results of the Science Analysis Software on the validated observations and containing mainly the high-level stellar parameters. Details are given in Section 4.
- A calibrator entry contains the following fields:
    - Calibrator ID (s)
    - RA, DEC
    - Magnitude V, R, J, H, K
    - Estimated angular diameter (UD_R, UD_H)
    - Flag for Primary Calibrators
- An observing block entry contains the following fields, organized in three levels:

- Description of the observation
  - OB number
  - PI Name and Program Name
  - A series of field containing the L0 metadata like: Night, Time, setup of the array, setup of the instrument,…
  - OB's type: Target/Calibrator/…
  - Target or Calibrator ID
- DRS results
  - A series of field containing the information on the result of the automatic data reduction software (including link to the raw oifits file)
- QCS results
  - A series of field containing the information on the result of the Quality Control Software (including link to the transfer function, to the calibrated oifits file if a target, ….)

3.2 Feeding the database

We distinguish six levels of access allowed to feed the database:

- Initial feeding (Arrow 1). The plan is to upload different lists of stars, that are currently being prepared by the Science Group. These lists are created through python codes querying different stellar databases, completing the information by different flags or computed parameters, and finally adding the required observing parameters. These python codes can easily provide cvs files or any other convenient format. It should be noted that this list of stars could be updated during the execution of the survey, as well as some flags of existing entries. This action corresponds to any update or initial feeding of any entries of the tables Targets or Calibrators. Only 'Manager' can do that.
- Entries following time allocation (Arrow 1). After the time allocation process and in advance to the night/run preparation, the users will be requested to provide the relevant information for the scheduling of their observations, in terms of configuration, number of measurements, and quality criteria. 'User' accounts are granted to this access. This action corresponds to any update or initial feeding of any entries of the table Targets.
- New entries related to the night scheduling (Arrow 4). If the night preparation procedure (Night Scheduling Software – NSS) identifies specific calibrators not yet in the database, new entries should be created. This action on the table Calibrators is granted to a Manager account only.
- Automatic SPICA-DRS output (Arrow 5). The SPICA-DRS (Data Reduction Software) will feed the database with the executed OBs (Table OB) and the corresponding metadata. The products of the DRS (raw oifits file) will be indexed in the database, and linked to the corresponding OB. The rights 'Manager' will be given to SPICA-DRS.
- The SPICA Quality Control Software (Arrow 7). This tool is dedicated to the determination of the transfer function of each night and to the estimation of the calibrated interferometric products. Quality controls are based on the science specifications (Quality criteria) contained in each Target entry. They permit to validate or not an observation. This high-level information is indexed in the database, as well as the transfer function of the night with the relevant information. For the Calibrators entries, only the transfer function will be indexed to the corresponding OB. Moreover, the completion rate factor will be updated if the OB is considered as successful. The rights 'Manager' will be given to SPICA-QCS.

- Science Analysis Software (Arrow 9). Based on the progress of the execution of the observations and the analysis of the QCS information, the SAS will upload the results of the high-level astrophysical analysis in the targets Table. 'Manager' and 'Users' can do that.

3.3 Querying the database

Five types of query are considered:

1. Consulting the execution of a night (Arrow 2). At any time, any registered user can check the execution of the nights with a presentation of the metadata, the transfer function and detailed information on any executed single observation. Any user granted with the correct access to the database or to some of its entries can download the high-level products of the DRS and QCS for further analysis. This corresponds to entries in the three tables.
2. Preparing a night of observation (Arrow 3). Through a client interfaced to the database, the Night Scheduling Software (NSS), can select (with the 'Runner' authorizations) the entries of the database (table Targets) validating a series of criteria. These entries are then exported and downloaded into the JMMC ASPRO2 software for the final preparation of the night. During the preparation, the table Calibrators is queried. If needed a 'Manager' account could update the calibrator Table.
3. Production of the calibrated products (Arrow 6). The QCS upload the results of the DRS of a night (Table OBs) for producing the calibrated quantities.
4. Production of Science-Ready data (Arrow 8). The SAS queries the database according to the rights granted to the account and extract all the relevant information (table OBs) necessary to produce the high-level science products.
5. Survey products (Arrow 10). This special action will permit a global analysis of the science-ready data for any SPICA proposals and for the survey analysis.

3.4 Description of the tools

3.4.1: SPICA Database Feeding Software (SPICA-DFS)

See slides

3.4.2 SPICA Night Scheduling Software

We consider the following sequence of operation:

- On the SPICA-DB interface, identification of the targets corresponding to a set of criteria (declination, magnitude, completion rate, priority…) through filters.
- Identification of a list of suitable calibrators, including at least N primary calibrators.
- Send these two lists to ASPRO2
- Use ASPRO2 to build the night strategy, including interrogation of the database to check for previous observations.
- Use the weighting function (priorities, completion rate, OBs, seeing, conditions…) to get the overall ranking of the OBs.
- When done, export the OBs as a single file for each OB and store them into the adequate folder of the SPICA-Observing Software machine.

3.4.3 SPICA Data Reduction Software

This software is executed automatically at the end of the night, after the sequence of archiving. The following steps are considered for each OB of the night:

- Load the data files and the associated auxillary data.
- Process the data
- Prepare the metadata describing the OB.
- Store the product of the DRS as a raw oifits file.
- Store the metadata and the raw oifits file on the SPICA-DRS machine.
- Create the corresponding OB in SPICA-DB and index all the related information.

3.4.4 SPICA Quality Control Software

When all the OBs of a night have been processed by SPICA-DRS, the SPICA-QCS will proceed with the following operations:

- Identify all the OB of the night corresponding to calibrators.
- Perform the computation of the transfer function for the whole night or for a number of segments, depending on the final configuration (instrument setup, number of telescopes).
- Store these transfer functions on the SPICA-DRS machine and index the corresponding OB in SPICA-DB.
- Perform the calibration of the raw oifits file of each OB corresponding to a target. Generate a calibrated oifits file and store it in the SPICA-DRS machine.
- Proceed to the control of the quality criteria and update the associated flags in the corresponding OB.
- Index the corresponding targets OB with the relevant information (transfer function, calibrated oifits file).

3.4.5 SPICA Science Analysis Software

- Select all the OB of a specific target, following the specific authorizations. If required the calibrated data will be exported for any specific processing.
- Depending on the type of observations (diameter or image), a model fitting will be done or an image reconstruction, respectively.
- A model fitting will be done using atmosphere models, like MARCS or ATLAS/Kurucz (etc...) depending on the spectral type and class of the object. The LD diameter and/or the LD coefficients will be derived.
- The SPICA-visibility analyzer will also check the departure to simple geometrical models and to standard atmosphere models to potentially flag unexpected stellar activity. The corresponding OB will be flagged. Possibly, the type of stellar activity will be identified, and new observations will be done with a different strategy.
- Extraction of radius (with Gaia parallaxes), effective temperature (with Fbol), mass and age in an automatic way. The purpose is to have a homogeneous set of fundamental parameters for about 800 stars all over the HD diagram. For the mass and age, grids of evolution models will be used, with an adaptation to the spectral type and class of the stars.
- For image, specific analysis will be done. Binaries, rotating stars, environment will have their own specific models and strategy.

- The images will be used to characterize stellar activity in order to better estimate the systematics on the SBCR, and possibly to correct the SBCR from these effects.
- All stars will have also specific outputs parameters as described by the different science groups. These parameters will be derived analytically and automatically in the SPICA-DB from astrophysical inputs parameters, DRS and QRS results. This will be possible mostly for stars with non-active angular diameter estimates. For other stars (significantly active or for stars with images), the calculation of output parameters will use specific numerical models.
- The final objective is to have homogeneous parameters (R, Teff, M, age) for about 1000 stars. Specific parameters, when the star is hosting a planet, asterosismic, pulsating, rotating, with a wind and/or environment or when it is in a binary, will also be extracted and stored in the database.
- The SPICA-DB data will be used for refereed publications, and codes will be used to extract the parameters and put them in a Vizier format in order to publish the various catalogues.

4. Database Specifications

As already mentioned, there are different parts in the DB that are connected: the "targets" and "calibrator" lists, and the Observing Blocks (OBs) list. The definition of the fields contained in these three tables is provided at the following link:

https://docs.google.com/spreadsheets/d/1EY3RvBAqsiv589YJn-8A_Tpmg_ORAdiYajRmS7Xc_-0/edit#gid=0