Solar Orbiter ARchive (SOAR) Status

Helen Rose Middleton

Solar Orbiter SOWG

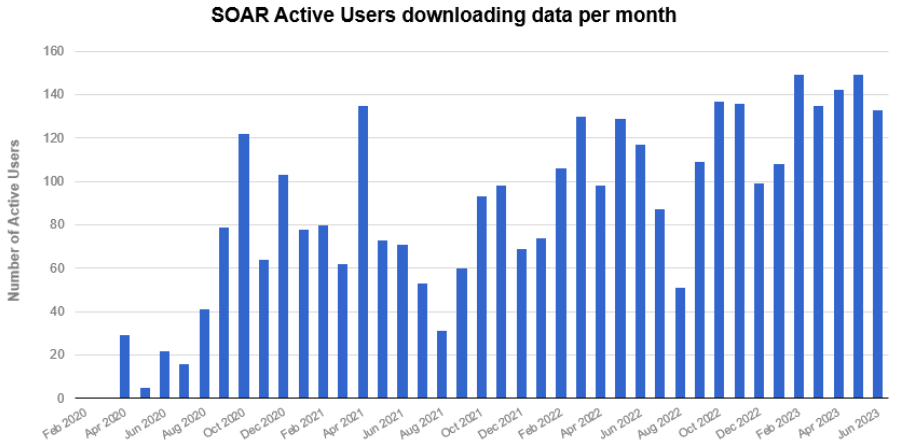
11/07/2023

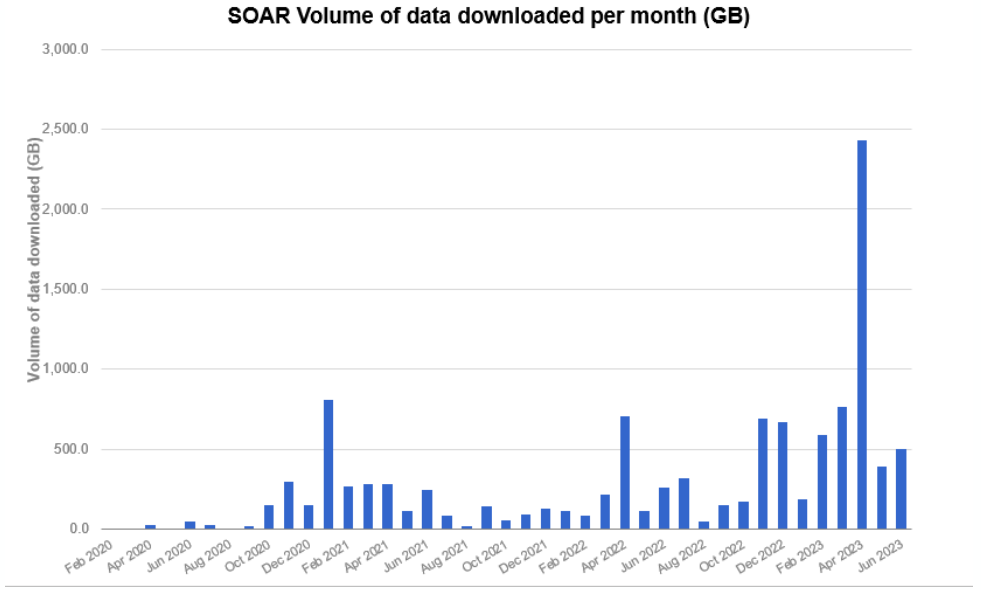
ESA UNCLASSIFIED – For ESA Official Use Only

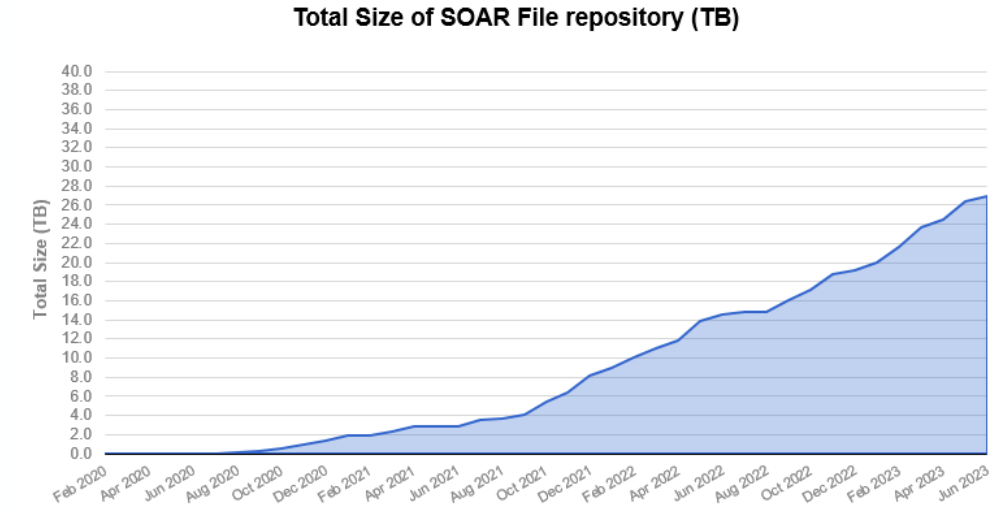
Contents

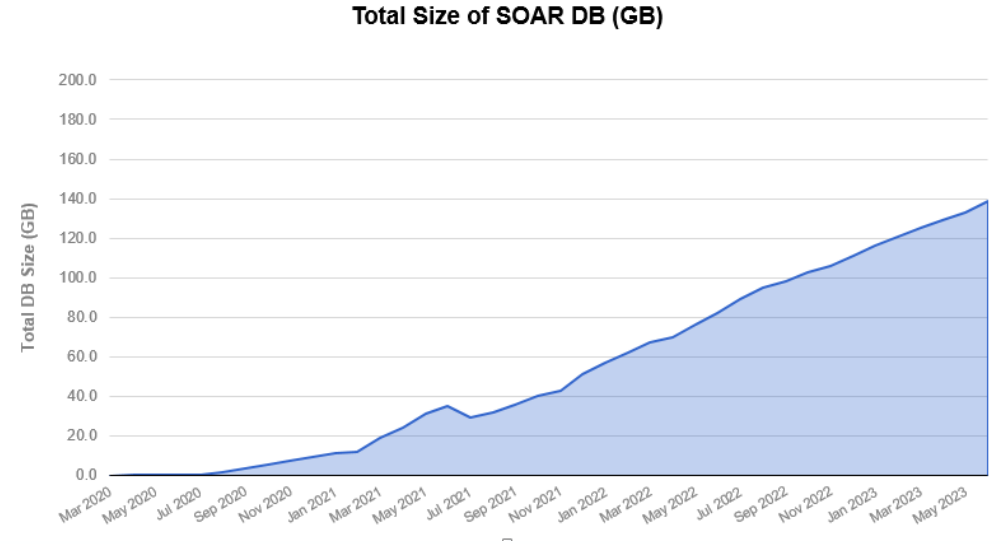
- Overall status
- new Beta:
    - Asynchronous download
    - Updated TAP tables allowing more flexible data requests
    - Deactivated files – searching on GUI and TAP
    - JP2 files – correct end times
    - Lower case descriptors
- Metadata Standard
    - HTTP_LINK and DOIs
    - Matters to discuss (light-time correction and DOIs)
    - State of SOOP metadata
- News on direct data transfer
- Datalabs
- Help pages

Active users, downloads and size of archive

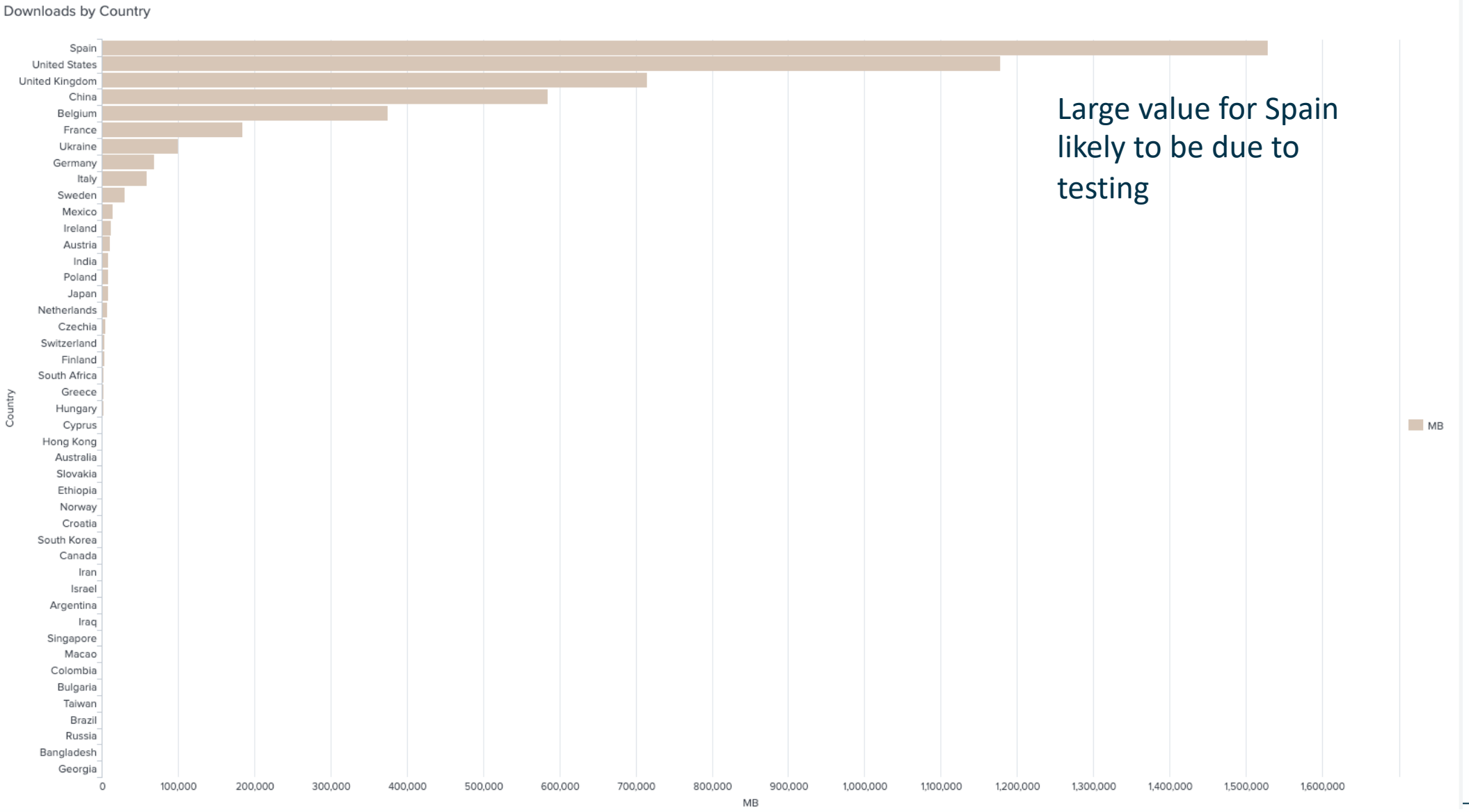








Downloads by Country



Large value for Spain likely to be due to testing

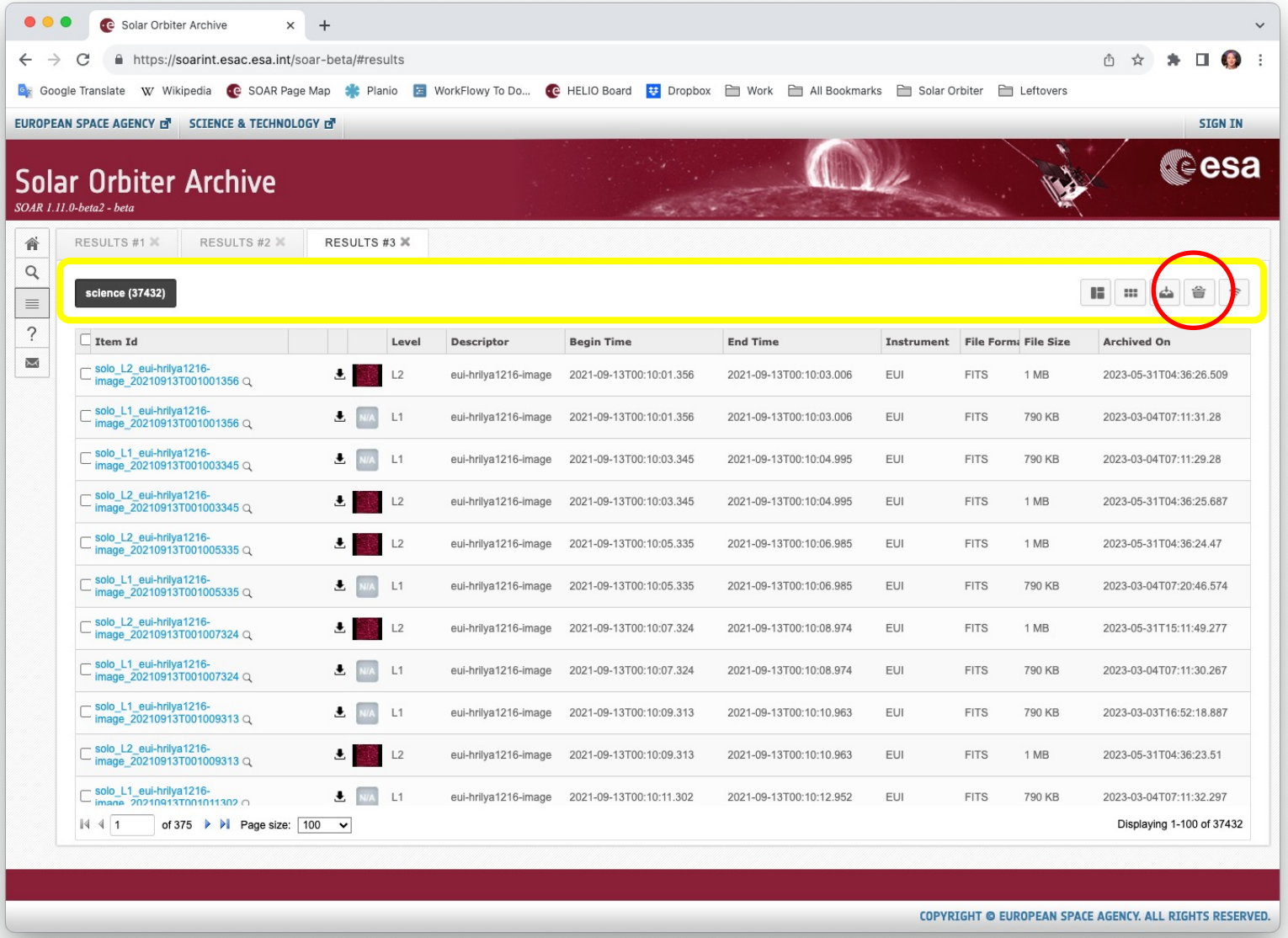
Beta features

SOAR Release 1.11.0

- Enhanced data download options
  - new asynchronous data download button
  - data download up to 50GB
  - removed 1000 file download limitation
  - enable TAP conditional downloads from data items (latest version not only all versions)
  - enhanced user interface notifications for download limits
- Upgrade to TAP 9.1.0
- Support searching for active/inactive science products (see next slide)
- Process start and end dates from JP2 files metadata
- Conversion of CDF and FITS descriptors to lower case

Should be released in the next week or so.

Asynchronous Download



For large requests under 50 GB, data will be packaged and an email sent when it is ready to collect.

Good for avoiding timeouts since it can take a while to assemble large requests.

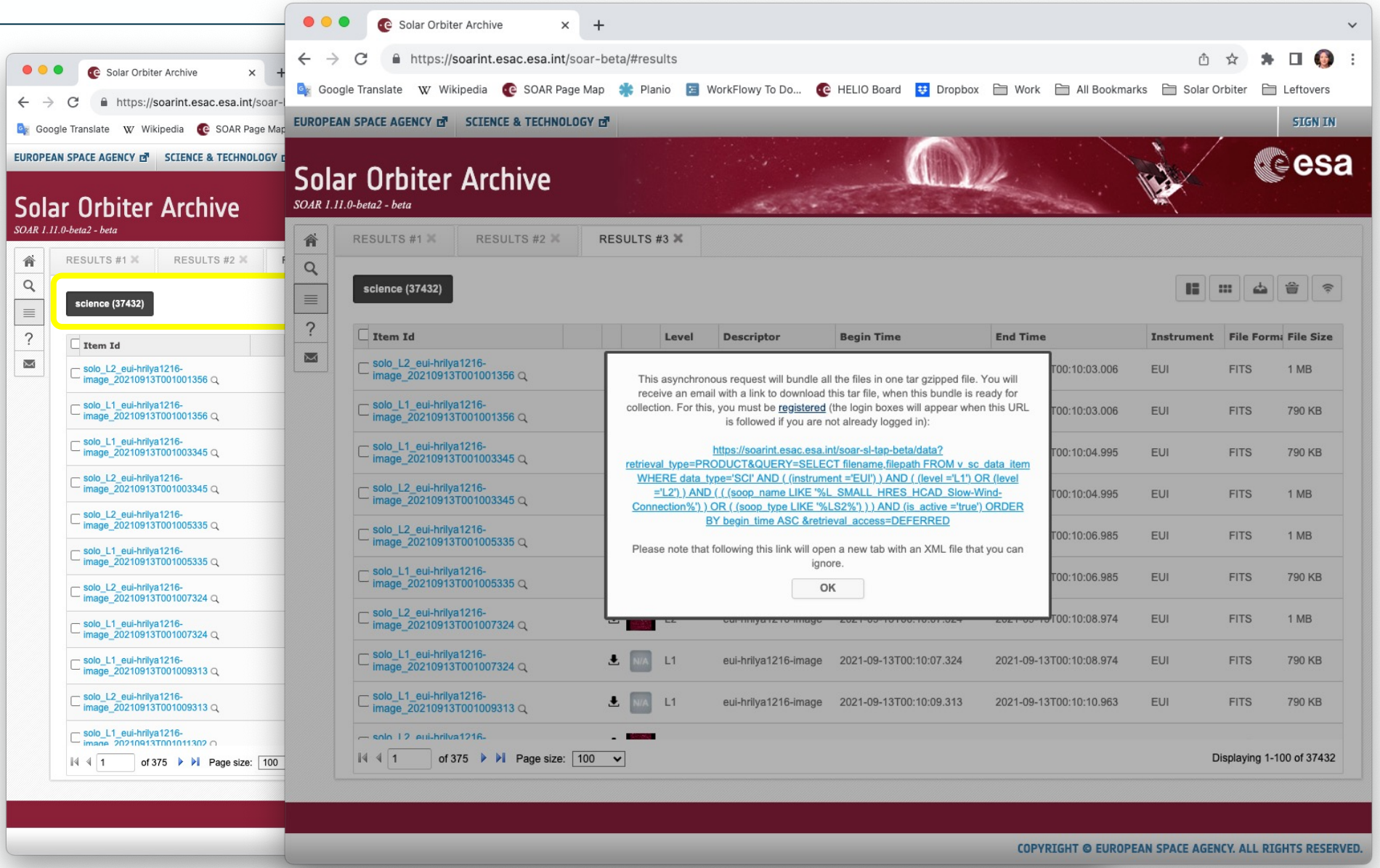
Async will only work for all items, and will grey out if any products are selected.

Asynchronous Download

The async button will actually provide the command line request for asynchronously downloading the results of the search that was done on the GUI.

Registration is necessary for asynchronous downloads – follow the link if not already registered.

If already registered, click the large link to start the request, or click OK to remove the message window.



Asynchronous Download

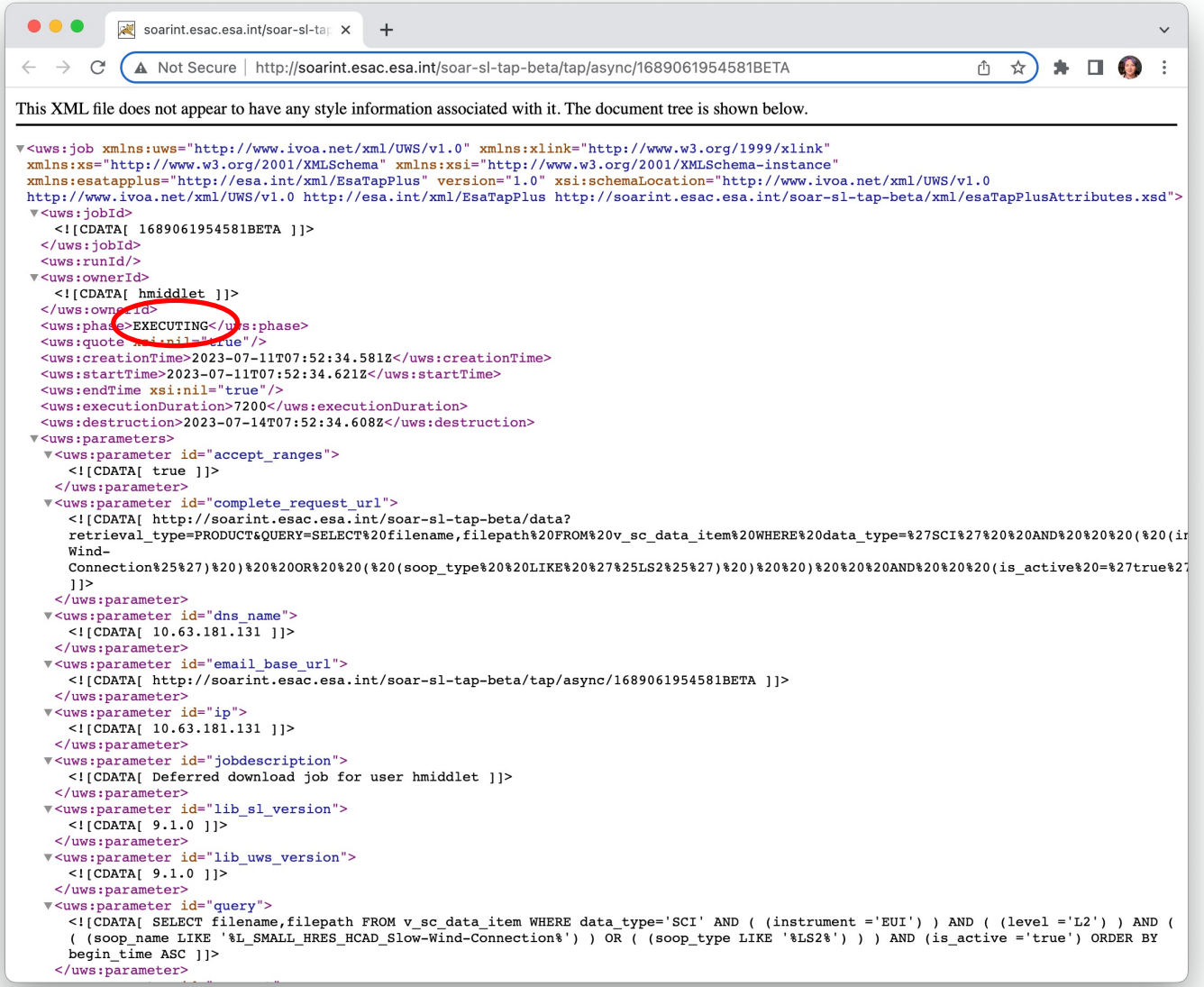
Once you've made the request (followed the long link), a new tab will open and ask you to log in.

When the request is successfully started, an XML page will appear with all the information associated with the request, including its current status – see red circle:

This is the response and can be used to automate the queries – see

https://www.cosmos.esa.int/web/soar/asynchronousrequests

for all the details!



Updated TAP Tables

While TAP metadata requests are very powerful, TAP data requests were more restrictive and were for only a few files at once because the individual data_item_id's were needed:

https://soar.esac.esa.int/soar-sl-tap/data? + …

* Either latest versions of individual files via data_item_id

retrieval_type=PRODUCT&data_item_id=solo_L2_epd-ept-north-rates_20200820,solo_L2_epd-ept-south-rates_20200820&product_type=SCIENCE

* Or latest low latency

retrieval_type=PRODUCTS&data_item_id= 'solo_LL02_epd-ept-north-rates_20200414T000009-20200415T000008, solo_LL02_epd-ept-south-rates_20200414T000009-20200415T000008&product_type=LOW_LATENCY

* Or all versions

retrieval_type=ALL_PRODUCTS&data_item_id=solo_L2_epd-ept-north-rates_20200820,solo_L2_epd-ept-south-rates_20200820&product_type=SCIENCE

TAP requests for more files

A metadata request can contain wildcards:

http://soar.esac.esa.int/soar-sl-tap/tap/sync?REQUEST=doQuery&LANG=ADQL&FORMAT=CSV&QUERY=SELECT+COUNT+(filename)+FROM+v_sc_data_item+WHERE+soop_name+LIKE+'%25CC_OFFPOI_FLATFIELD_HRI%25'

To be able to use wildcards (% and _, equivalent to more usual * and ?) in a data request, the search is for **filepath** and **filename**:

http://soar.esac.esa.int/soar-sl-tap/data?retrieval_type=PRODUCT&QUERY=SELECT+filepath,filename+FROM+v_sc_repository_file+WHERE+soop_name+LIKE+'%25CC_OFFPOI_FLATFIELD_HRI%25'

Filename and filepath weren't available in the table for the latest data, v_sc_data_item (or v_ll_data_item) so before the Beta v1.11, large wildcard downloads were possible for table containing all versions (v_sc_repository_file), but not the from the latest version table (v_sc_data_item).

TAP requests for more files

A metadata request can contain wildcards:

http://soar.esac.esa.int/soar-sl-tap/tap/sync?REQUEST=doQuery&LANG=ADQL&FORMAT=CSV&QUERY=SELECT+COUNT+(filename)+FROM+v_sc_data_item+WHERE+soop_name+LIKE+'%25CC_OFFPOI_FLATFIELD_HRI%25’

To be able to use wildcards (% and _, equivalent to more usual * and ?) in a data request, the search is for **filepath** and **filename**:

http://soar.esac.esa.int/soar-sl-tap/data?retrieval_type=PRODUCT&QUERY=SELECT+filepath,filename+FROM+V_SC_DATA_ITEM+WHERE+soop_name+LIKE+'%25CC_OFFPOI_FLATFIELD_HRI%25’

Filename and filepath weren’t available in the table for the latest data, v_sc_data_item (or v_ll_data_item) so before the Beta v1.11, large wildcard downloads were possible for table containing all versions (v_sc_repository_file), but not the from the latest version table (v_sc_data_item). BUT NOW IT’S POSSIBLE!

Deactivations

Background: EUI wanted to change some of their descriptors, to make them more specific:

So, for example, for the files with descriptor *eui-fsi174-image*, after Data Release 6, 1782 + 9544 new versions of the same observations would now have different names:

eui-fsi174-image	104636
eui-fsi174-image-occulter	1782
eui-fsi174-image-short	9544

Without deactivating the files previous to this Data Release, a file with the new name *eui-fsi174-image-occulter* would show up against its old version's name *eui-fsi174-image* and that latest version of the old name would always appear in the search.

Deactivations

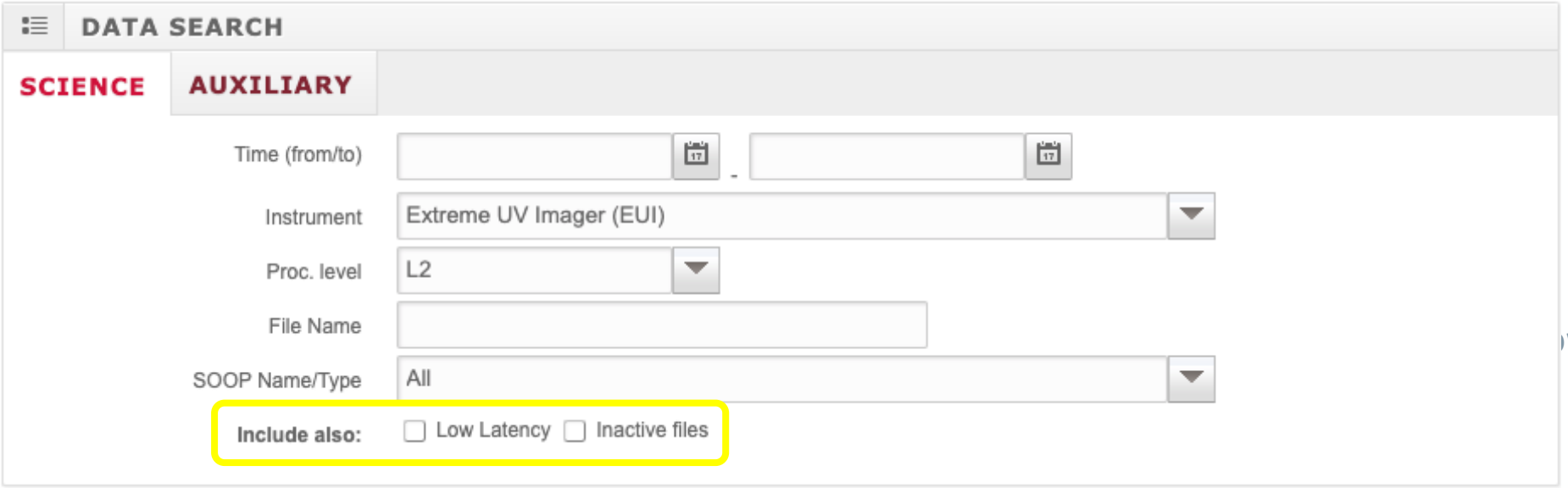
EUI wanted to change some of their descriptors, to make them more specific:

So, for example, for the files with descriptor *eui-fsi174-image*, after Data Release 6, 1782 + 9544 new versions of the same observations would now have different names:

eui-fsi174-image

eui-fsi174-image-occu

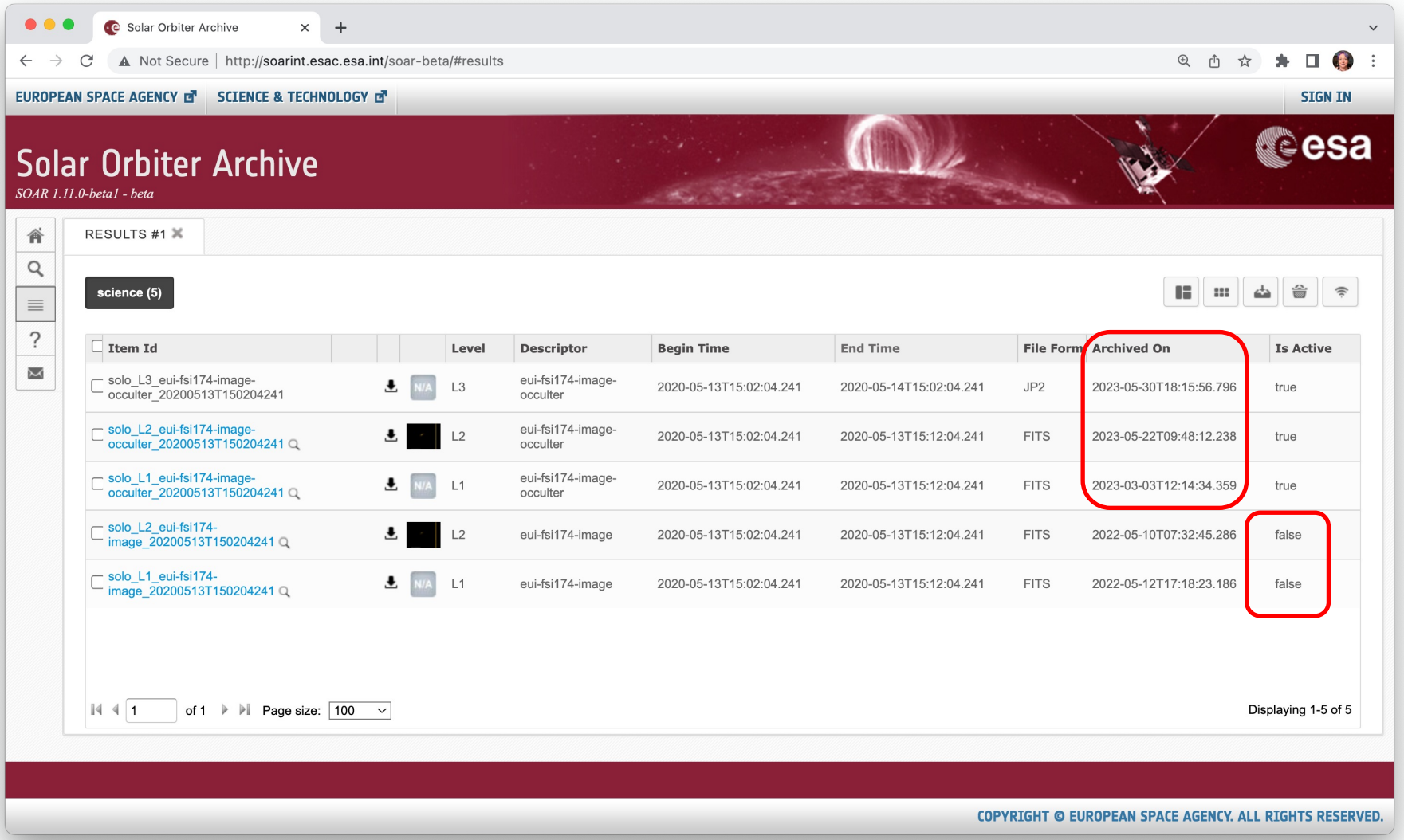
eui-fsi174-image-shor



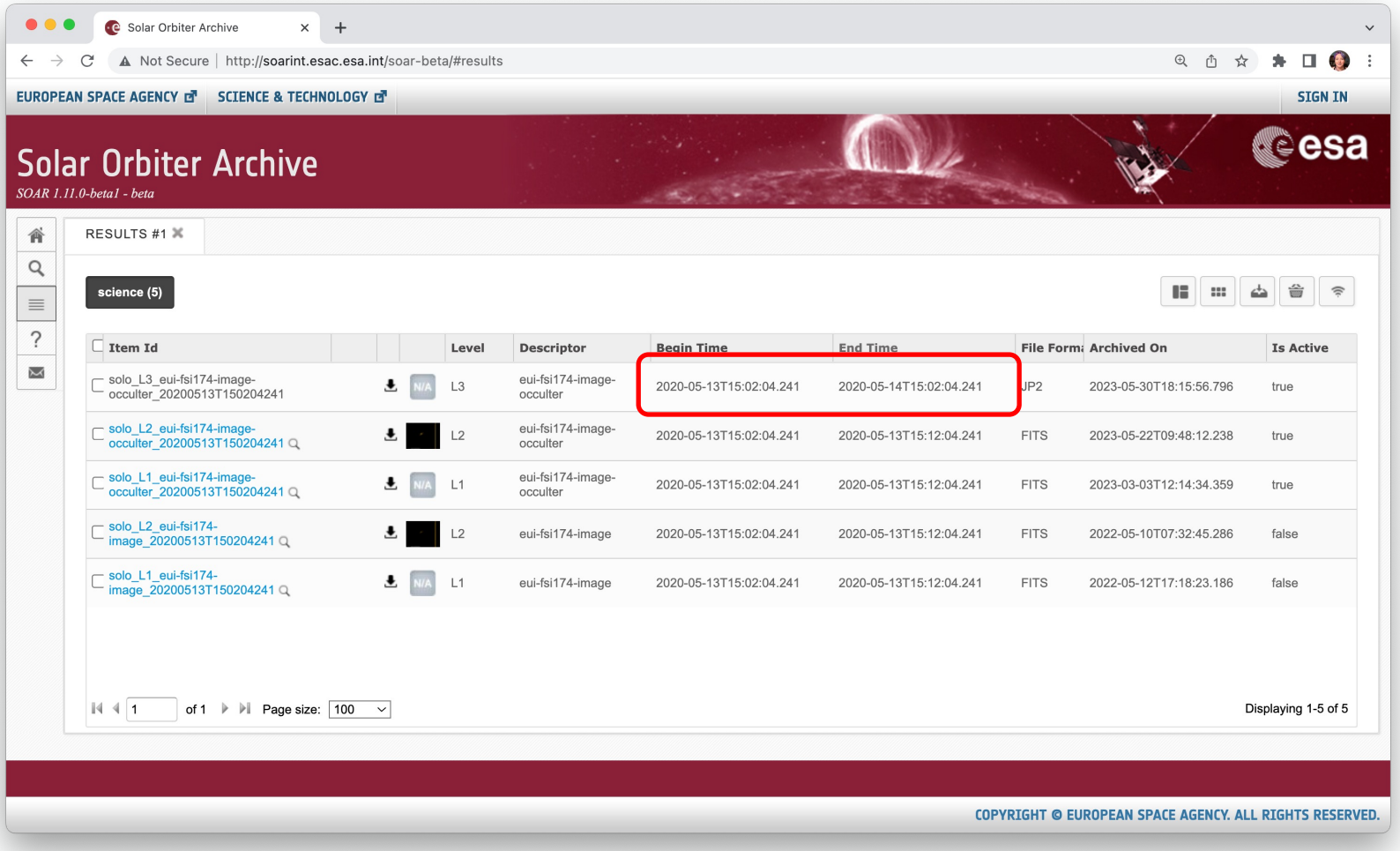
Without deactivating the fil ow up against its old version's r search.

+ New version names included the # symbol, e.g., eui-hrieuv###-image. We briefly allowed, but causes problems for downloading and mirroring the SOAR, so no longer allowed and EUI will change to xxx

Inactive files



JP2 metadata parsed for times (not re-ingested yet)



The JP2 files weren't parsed for the metadata, so because of the default 24-hour length file if the end date isn't included in the filename, the end times were wrong.

This has been fixed and the JP2 metadata is now parsed to give the correct end times.

They will all need to be re-ingested for the new information to get into the database

Lower-case descriptors

science (5)

Item Id			Level	Descriptor	Begin Time
solo_L3_eui-fsi174-image-occulter_20200513T150204241		N/A	L3	eui-fsi174-image-occulter	2020-05-13T15:02:04
solo_L2_eui-fsi174-image-occulter_20200513T150204241			L2	EUI-FSI174-IMAGE-OCCULTER	2020-05-13T15:02:04
solo_L1_eui-fsi174-image-occulter_20200513T150204241		N/A	L1	EUI-FSI174-IMAGE-OCCULTER	2020-05-13T15:02:04
solo_L2_eui-fsi174-image_20200513T150204241			L2	EUI-FSI174-IMAGE	2020-05-13T15:02:04
solo_L1_eui-fsi174-image_20200513T150204241		N/A	L1	EUI-FSI174-IMAGE	2020-05-13T15:02:04

The descriptors were provided in a mixture of upper case and lower case.

Lower-case descriptors

science (5)

Item Id			Level	Descriptor	Begin Time
solo_L3_eui-fsi174-image-occulter_20200513T150204241		N/A	L3	eui-fsi174-image-occulter	2020-05-13T15:0
solo_L2_eui-fsi174-image-occulter_20200513T150204241			L2	eui-fsi174-image-occulter	2020-05-13T15:0
solo_L1_eui-fsi174-image-occulter_20200513T150204241		N/A	L1	eui-fsi174-image-occulter	2020-05-13T15:0
solo_L2_eui-fsi174-image_20200513T150204241			L2	eui-fsi174-image	2020-05-13T15:0
solo_L1_eui-fsi174-image_20200513T150204241		N/A	L1	eui-fsi174-image	2020-05-13T15:0

The descriptors were provided in a mixture of upper case and lower case.

This has been corrected so that they are all provided in lower case.

Lower-case descriptors

science (5)

Item Id			Level	Descriptor	Begin Time
solo_L3_eui-fsi174-image-occulter_20200513T150204241		N/A	L3	eui-fsi174-image-occulter	2020-05-13T15:0
solo_L2_eui-fsi174-image-occulter_20200513T150204241			L2	eui-fsi174-image-occulter	2020-05-13T15:0
solo_L1_eui-fsi174-image-occulter_20200513T150204241		N/A	L1	eui-f occu	
solo_L2_eui-fsi174-image_20200513T150204241			L2	eui-f	
solo_L1_eui-fsi174-image_20200513T150204241		N/A	L1	eui-f	

The descriptors were provided in a mixture of upper case and lower case.

This has been corrected so that they are all provided in

SWA-EAS filenames we've agreed with SWA to change:

<u>Logical Source and number of files</u>
solo_L1_swa-eas-OnbPartMoms_ 83
solo_L1_swa-eas1-NM3D_ 1932
solo_L1_swa-eas1-SSc_ 634
solo_L1_swa-eas2-NM3D_ 1866
solo_L1_swa-eas2-SSc_ 654 -> 654

plus the Metadata Standard will be updated.

Metadata Standard

v2.5 released in December 2022

Clarifications of sections making up the file naming convention, use of the BLANK keyword, acceptable ISTP keyword values (CDF)

Now mandatory: PARENT (FITS), Parents (CDF) made mandatory, CDF keyword SOOP_NAME, OBS_MODE (describe the mode in a meaningful way)

Now optional: TARGET, TDIMn for BINTABLE extension

For SOOP/OBS keywords (FITS and CDF) any not associated with SOOPs or OBS_ID should read **'none'**

Notes added to explain the structure of the OBT (FITS) and SCET (CDF) as coarse.fine

Section added for compressed FITS

v2.6 to be released

Minor changes including language about being a proposal: in this mature state it provides **requirements**

HTTP_LINK will be made optional, since not everyone hosts the data elsewhere (see next slide)

Section 3.1.1.8 on light-time correction?

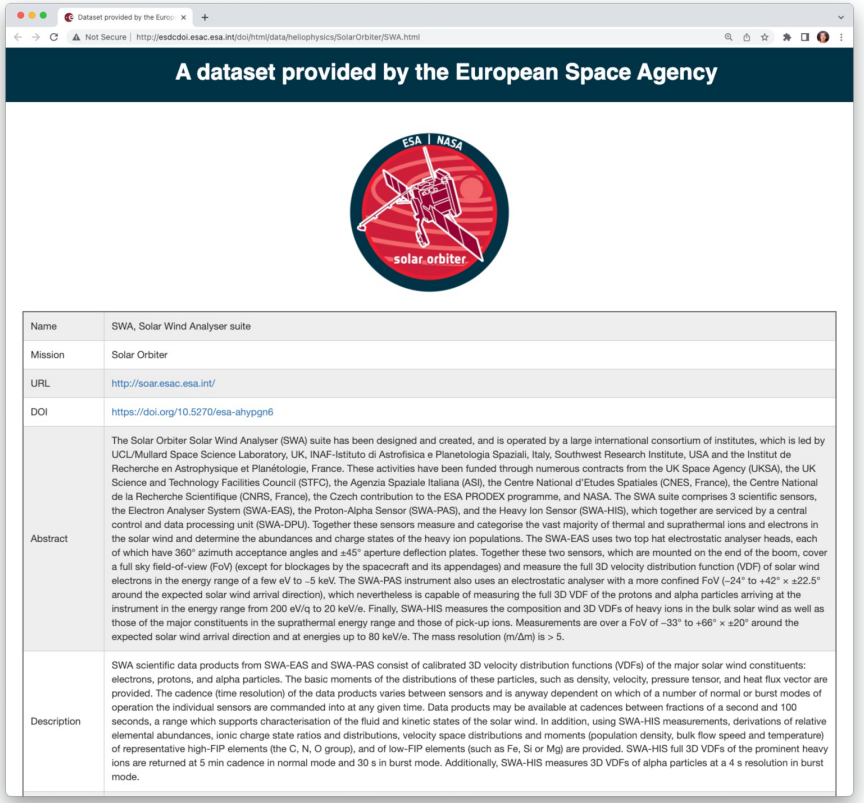
HTTP_LINK and DOIs

- HTTP_LINK was mandatory for CDFs, but recommended to be “The URL for the PI or Co-I web site holding online data.”
- SWA-HIS only host on SOAR, but do have usage notes.

EPD: https://doi.org/10.5270/esa-5897yve
EUI: https://doi.org/10.5270/esa-l169c5q
MAG: https://doi.org/10.5270/esa-ux7y320
Metis: https://doi.org/10.5270/esa-366ut35
SPICE: https://doi.org/10.5270/esa-lbmdy7c
SWA: https://doi.org/10.5270/esa-ahypgn6

Six instrument already have DOIs, list at https://www.cosmos.esa.int/web/esdc/doi/solarorbiter, for those who responded to the call from Arnaud Masson in March 2021

For SWA-HIS, proposed to add support data link and description to SWA DOI landing page (which can be updated), and include the SWA DOI in the TEXT field of the files with a short explanation.



Other updates to support data

STIX Re-requests

Some STIX observations have been re-requested from the spacecraft to adjust the resolution (spectral & temporal) with respect to the originally downlinked file. These re-requests are therefore different versions of already existing products, and so are given an additional '-sup1' or -sup2' to the descriptor. The original files and reasons for the re-request are given in this CSV file: supplement_report_20230517.csv

PHI Data Release Information

DPDD: SOL-PHI-MPS-SW7400-IF-2_0_1.pdf

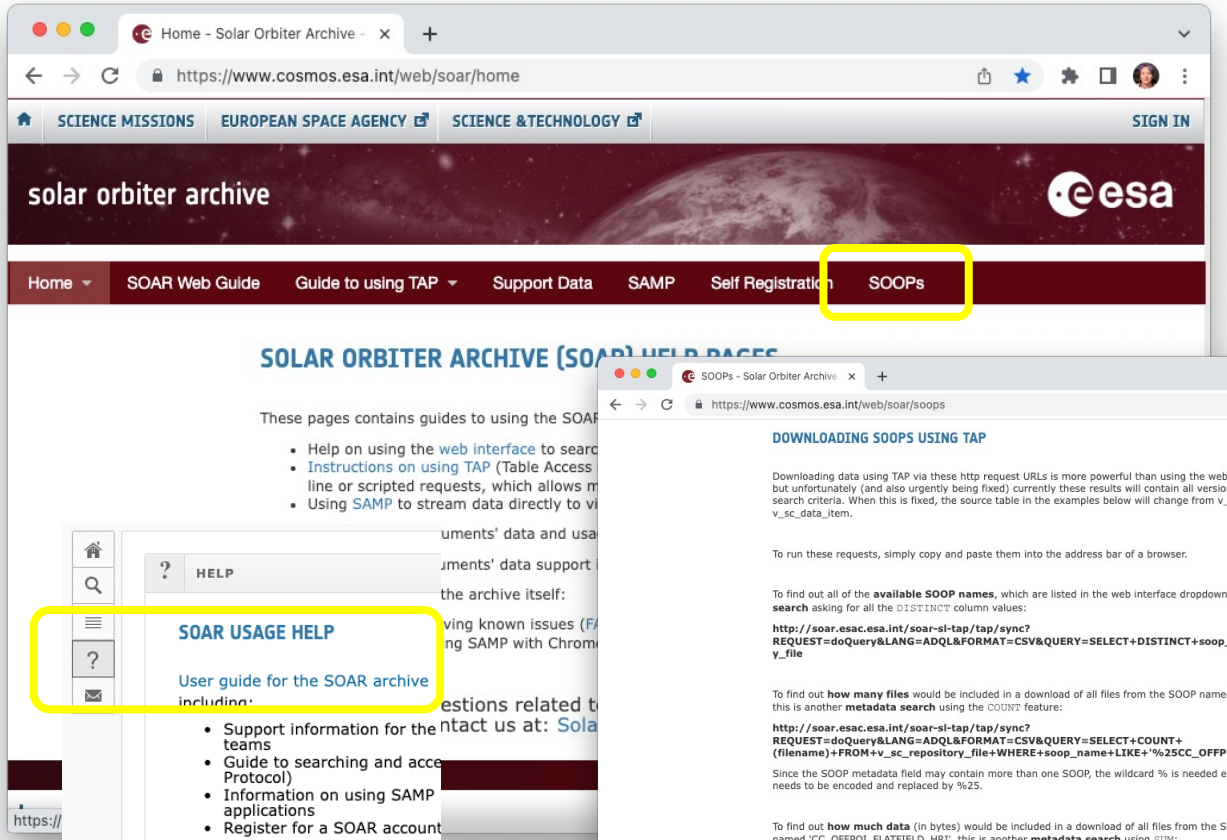
SO_PHI-HRT second data release.pdf

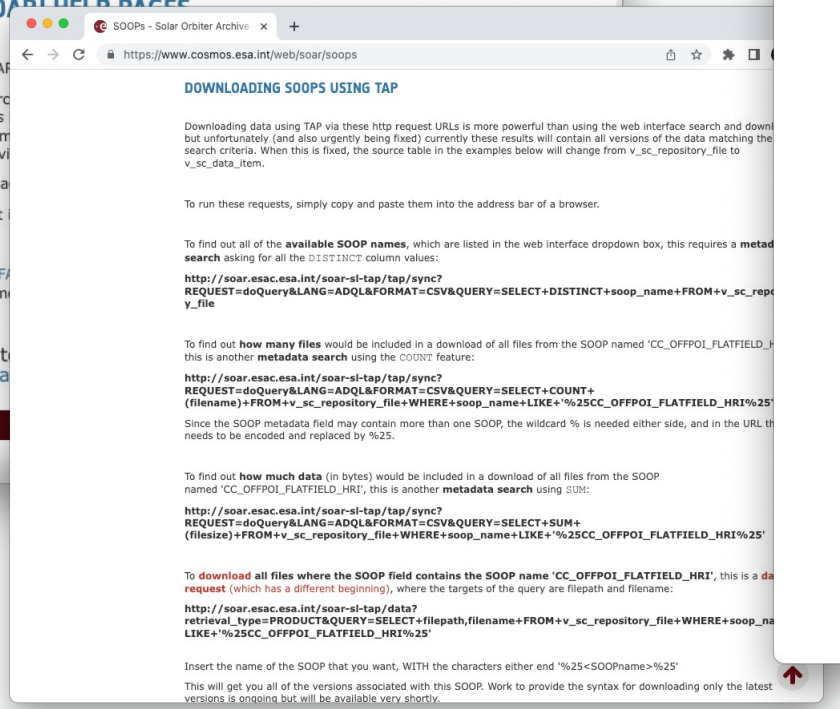
These are all given on the Cosmos SOAR Help Guide pages, along with the TAP help, Jupyter Notebooks and SOOP information:

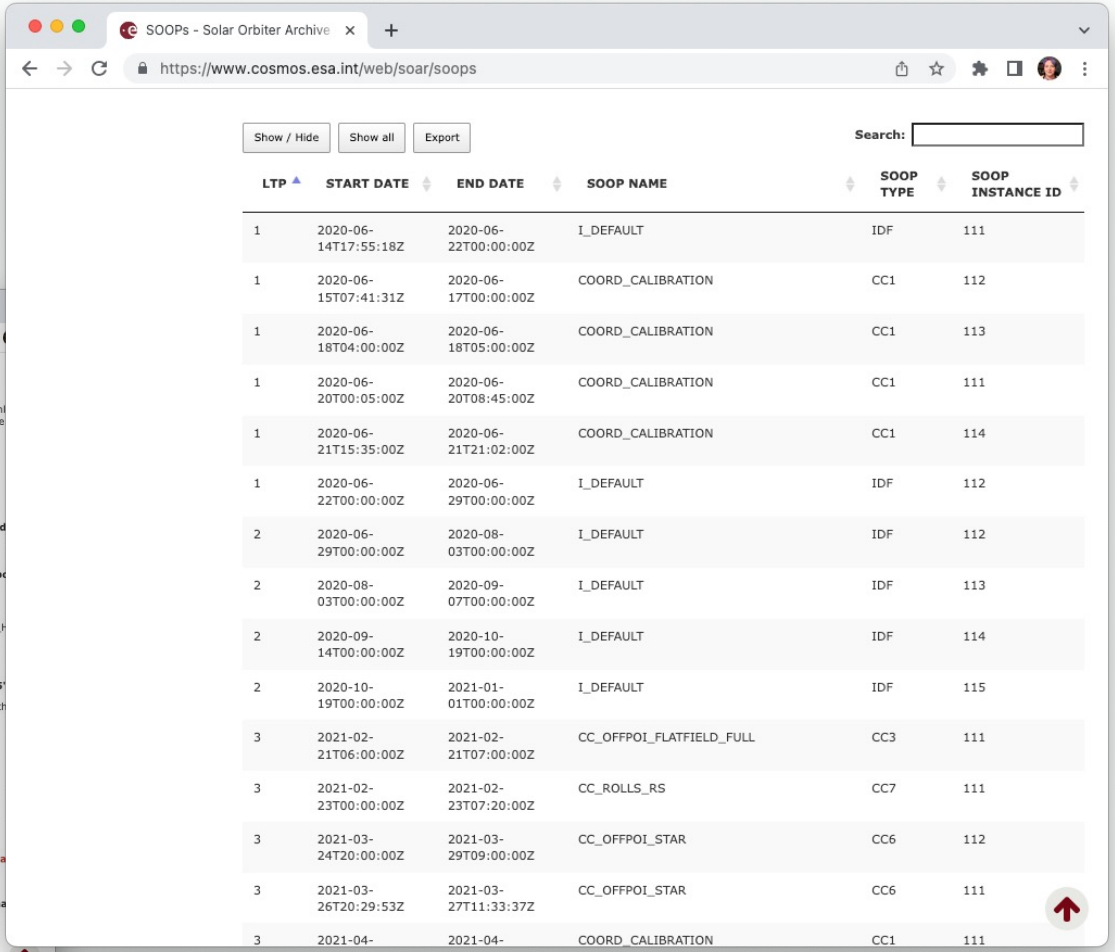
https://www.cosmos.esa.int/web/soar/home

Search by SOOP

If a SOOP isn't specified, intervals are listed in Help Pages: https://www.cosmos.esa.int/web/soar/soops



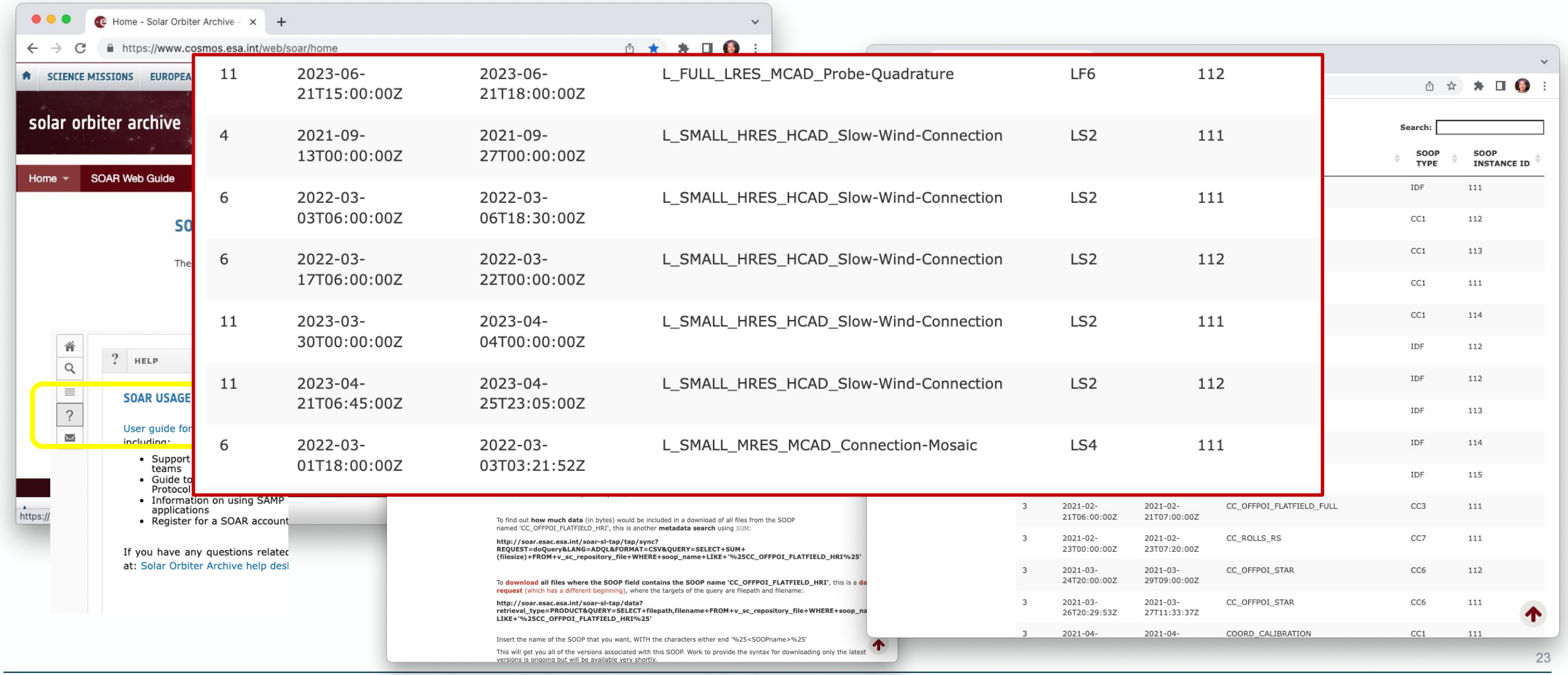




LTP	START DATE	END DATE	SOOP NAME	SOOP TYPE	SOOP INSTANCE ID
1	2020-06-14T17:55:18Z	2020-06-22T00:00:00Z	I_DEFAULT	IDF	111
1	2020-06-15T07:41:31Z	2020-06-17T00:00:00Z	COORD_CALIBRATION	CC1	112
1	2020-06-18T04:00:00Z	2020-06-18T05:00:00Z	COORD_CALIBRATION	CC1	113
1	2020-06-20T00:05:00Z	2020-06-20T08:45:00Z	COORD_CALIBRATION	CC1	111
1	2020-06-21T15:35:00Z	2020-06-21T21:02:00Z	COORD_CALIBRATION	CC1	114
1	2020-06-22T00:00:00Z	2020-06-29T00:00:00Z	I_DEFAULT	IDF	112
2	2020-06-29T00:00:00Z	2020-08-03T00:00:00Z	I_DEFAULT	IDF	112
2	2020-08-03T00:00:00Z	2020-09-07T00:00:00Z	I_DEFAULT	IDF	113
2	2020-09-14T00:00:00Z	2020-10-19T00:00:00Z	I_DEFAULT	IDF	114
2	2020-10-19T00:00:00Z	2021-01-01T00:00:00Z	I_DEFAULT	IDF	115
3	2021-02-21T06:00:00Z	2021-02-21T07:00:00Z	CC_OFFPOI_FLATFIELD_FULL	CC3	111
3	2021-02-23T00:00:00Z	2021-02-23T07:20:00Z	CC_ROLLS_RS	CC7	111
3	2021-03-24T20:00:00Z	2021-03-29T09:00:00Z	CC_OFFPOI_STAR	CC6	112
3	2021-03-26T20:29:53Z	2021-03-27T11:33:37Z	CC_OFFPOI_STAR	CC6	111
3	2021-04-	2021-04-	COORD_CALIBRATION	CC1	111

Search by SOOP

If a SOOP isn't specified, intervals are listed in Help Pages: https://www.cosmos.esa.int/web/soar/soops



11	2023-06-21T15:00:00Z	2023-06-21T18:00:00Z	L_FULL_LRES_MCAD_Probe-Quadrature	LF6	112
4	2021-09-13T00:00:00Z	2021-09-27T00:00:00Z	L_SMALL_HRES_HCAD_Slow-Wind-Connection	LS2	111
6	2022-03-03T06:00:00Z	2022-03-06T18:30:00Z	L_SMALL_HRES_HCAD_Slow-Wind-Connection	LS2	111
6	2022-03-17T06:00:00Z	2022-03-22T00:00:00Z	L_SMALL_HRES_HCAD_Slow-Wind-Connection	LS2	112
11	2023-03-30T00:00:00Z	2023-04-04T00:00:00Z	L_SMALL_HRES_HCAD_Slow-Wind-Connection	LS2	111
11	2023-04-21T06:45:00Z	2023-04-25T23:05:00Z	L_SMALL_HRES_HCAD_Slow-Wind-Connection	LS2	112
6	2022-03-01T18:00:00Z	2022-03-03T03:21:52Z	L_SMALL_MRES_MCAD_Connection-Mosaic	LS4	111

SOOP NAME

All fine except where not associated:

- ”
- unknown_soop
- not defined
- None
- **none**

Has soop_type but no soop_name: 42884

Has soop_name but no soop_type: 0

Has neither soop_type or soop_name: 397650

Has both soop_type and soop_name: 835167

SOOPs that were run with no mention in any metadata – maybe not received that data yet

LTP 6, 2022-01-20T00:00:00Z/2022-02-03T03:15:00Z, R_FULL_LRES_HCAD_Full-Disk-Helioseismology

LTP 11, 2023-04-05T09:00:00Z/2023-04-05T16:58:49Z, L_BOTH_HRES_HCAD_Major-Flare

SOOP Type for in situ

MAG: NULL, whitespace, IDF, NA

SOOP name in soop_type:

EPD: NULL, whitespace, I_DEFAULT (6,843), L_SMALL_HRES_HCAD_Slow-Wind-Connection (294)

RPW: NULL, whitespace, I_DEFAULT (6,194), L_SMALL_HRES_HCAD_Slow-Wind-Connection (252)

SWA-EAS: 149 different flavours of garbage ---------->

WHERE instrument!='SWA' (i.e., remote sensing): '' – 1,012

'000' – 128,119,

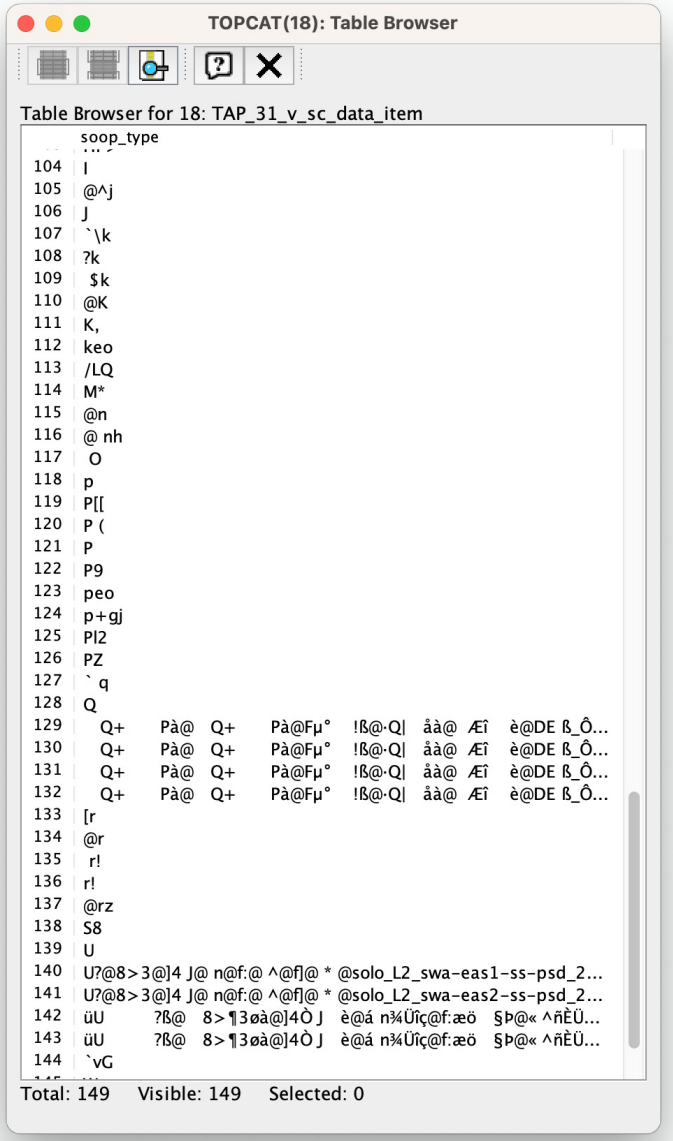
' ' – 4,227

' 000' - 255

STIX five files with 25 SOOPs

CC1;CC2;CC3;CC5;CC6;CC7;CC8;LB1;LB5;LF1;LF2;LF6;LS2;LS4;LS5;LTM;RB1;RB3;RF1;RS0;RS4;RS5;RS9;RSC;RSD

'none' 390,010 – thank you!

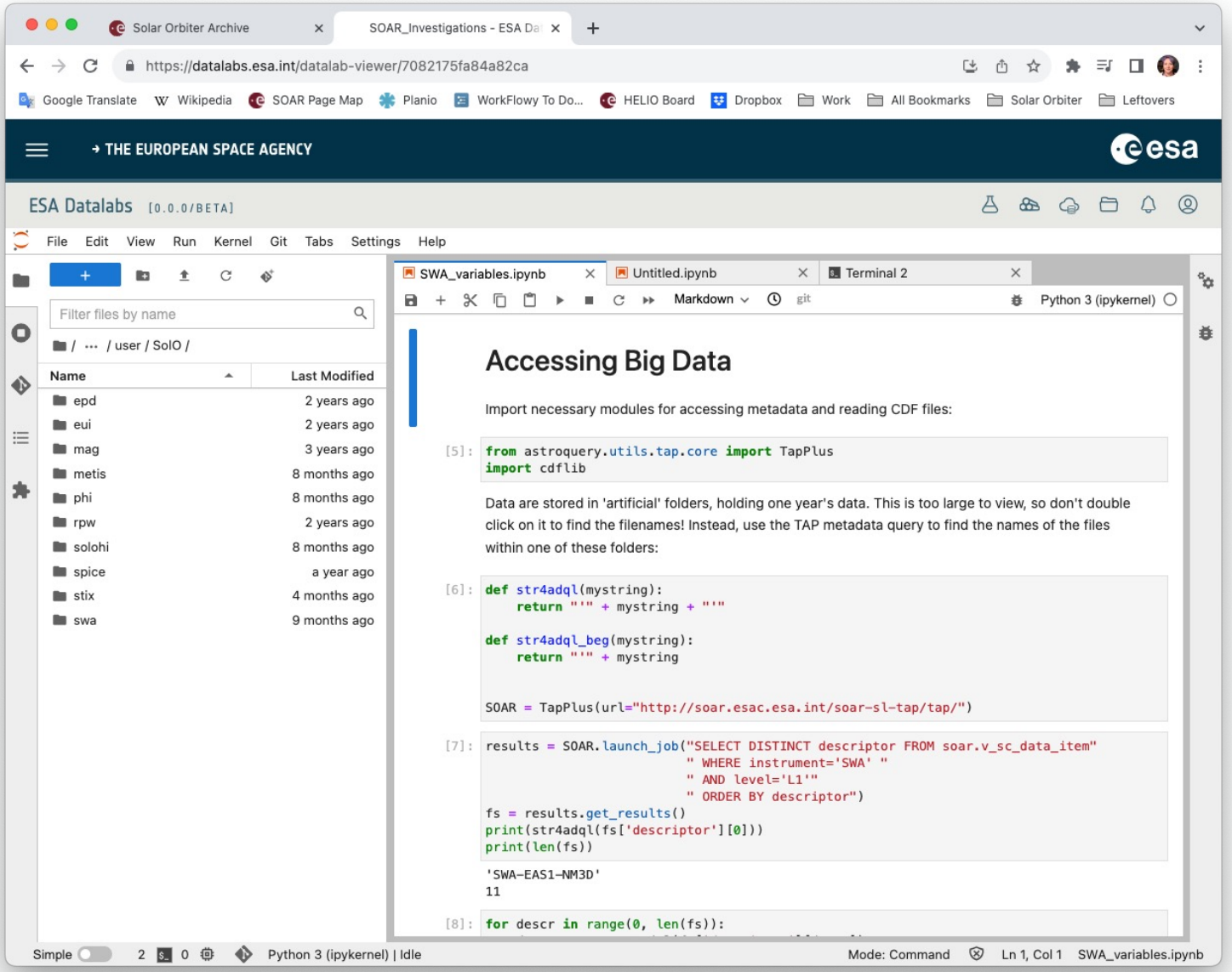


Direct upload from teams to SOAR

- EUI done
- ICD sent out to all teams
- Responses from SPICE and PHI
- SPICE has already done the required changes, so will be able to connect them this summer
- Other teams can be added in the autumn when more interaction is easier

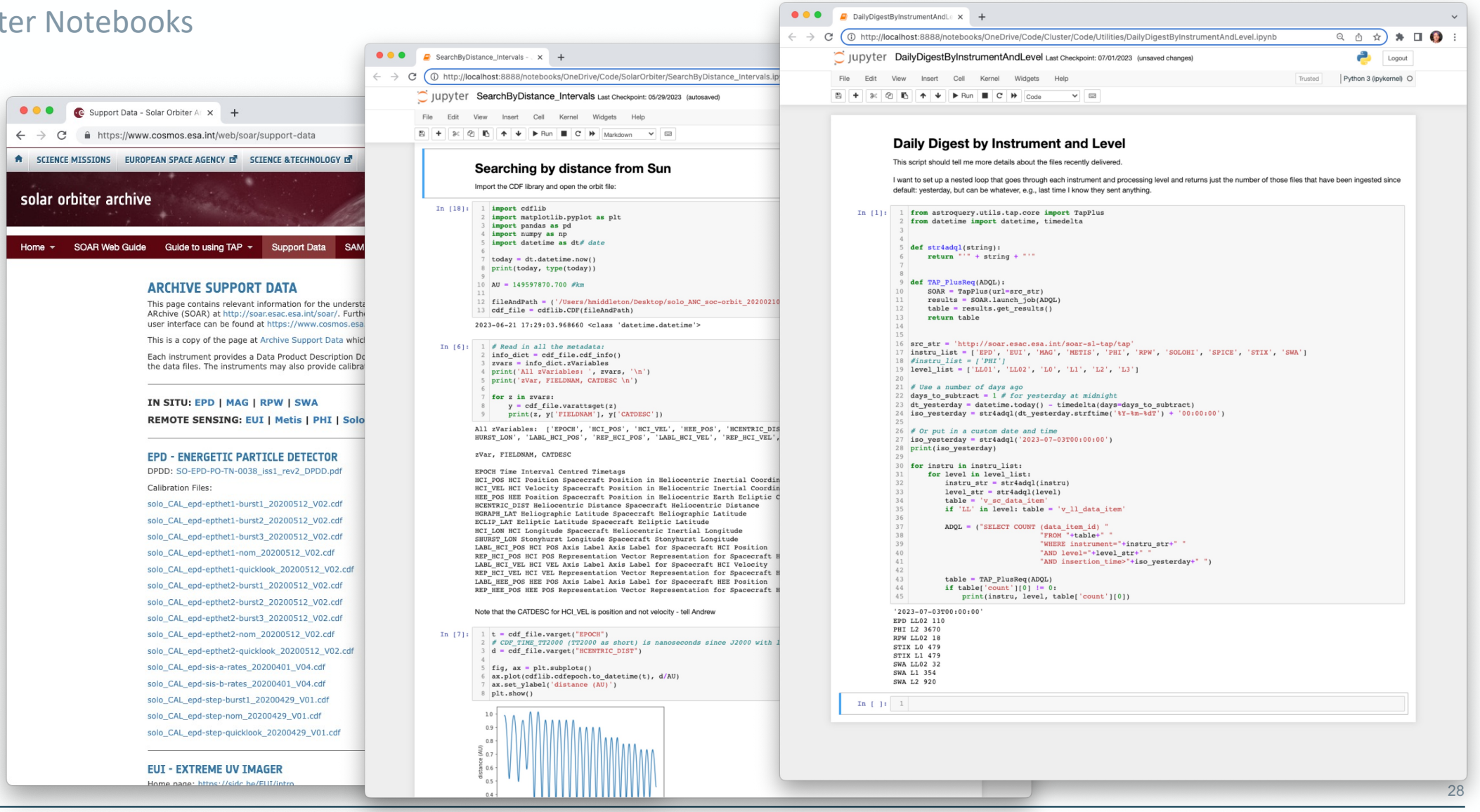
Datalabs

Datalabs provides Docker applications, including Jupyter Notebooks, alongside the entire SOAR, locally mounted for direct, complete and fast access to all archived data.



Help pages: https://www.cosmos.esa.int/web/soar/home

TAP help and Jupyter Notebooks



Me

- Helen Middleton

- I am 50% SOAR (along with 50% Cluster Science Archive), but if you have any doubts, questions, comments or requests for assistance, do not hesitate to contact me:

- Tel: +34 620 797 872
- E-mail: helen.middleton@esa.int