ESAC TAP Upload Use Cases

IVOA November 2024 Interoperability Meeting
Jose Osinde (Starion for ESA)
C. Rios (Starion for ESA, M. Henar (Starion for ESA), J. Ballester (Starion for ESA)
R. Parejo (Starion for ESA), R. Bhatawdekar (ESA)
SCO-08: Archives Software Development

ESAC
Camino Bajo del Castillo s/n, Urb. Villafranca Del Castillo
28692 Villanueva de la Cañada (Madrid) Spain

Data upload requirements in ESAC archives using TAP

The ESA TAP+ upload features provide end users with data upload options, enabling dynamic data input for both temporary and long-term storage needs. Two primary upload types, "On-the-fly" and "Persistent," serve distinct purposes:

- "On-the-fly" uploads support immediate, query-specific data that is deleted once the query is finished.

- "Persistent" uploads allow users to upload data via a local file or a URL, store it within a user-specific schema, and optionally share or reuse query results.

"On-the-fly" Upload Capabilities

- Part of the standard since TAP 1.0
- Immediate, query-specific data support: UPLOAD/QUERY/DELETE
- Tables uploaded to a custom schema and deleted once the query is finished
- Typical use cases:
  - Crossmatch of local sources against a table in the server
  - Select a list of rows in the server according to a list of elements available locally
  - Use local results with a remote TAP service

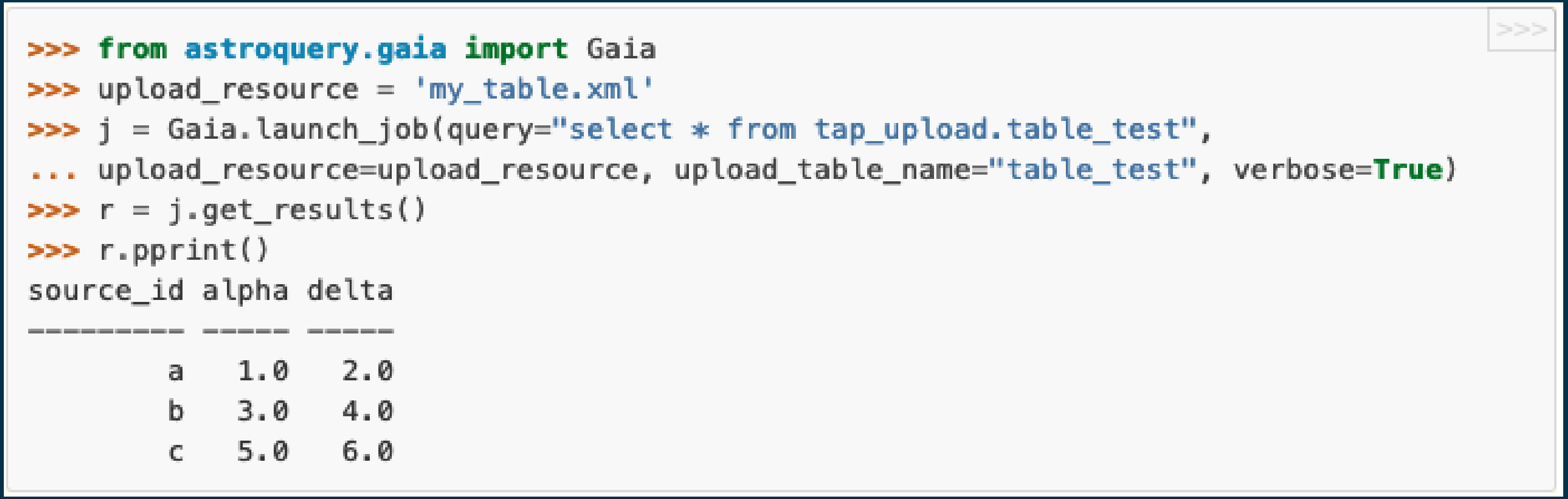
"On-the-fly" Upload Capabilities

- Gaia example (https://astroquery.readthedocs.io/en/latest/gaia/gaia.html):

1.5. Synchronous query on an 'on-the-fly' uploaded table

A votable can be uploaded to the server in order to be used in a query.

You have to provide the local path to the file you want to upload. In the following example, the file 'my_table.xml' is located to the relative location where your python program is running.



```
>>> from astroquery.gaia import Gaia
>>> upload_resource = 'my_table.xml'
>>> j = Gaia.launch_job(query="select * from tap_upload.table_test",
... upload_resource=upload_resource, upload_table_name="table_test", verbose=True)
>>> r = j.get_results()
>>> r.pprint()
source_id alpha delta
--------- ----- -----
        a   1.0   2.0
        b   3.0   4.0
        c   5.0   6.0
```

"On-the-fly" Upload Capabilities

- Fully supported by other applications as TopCat

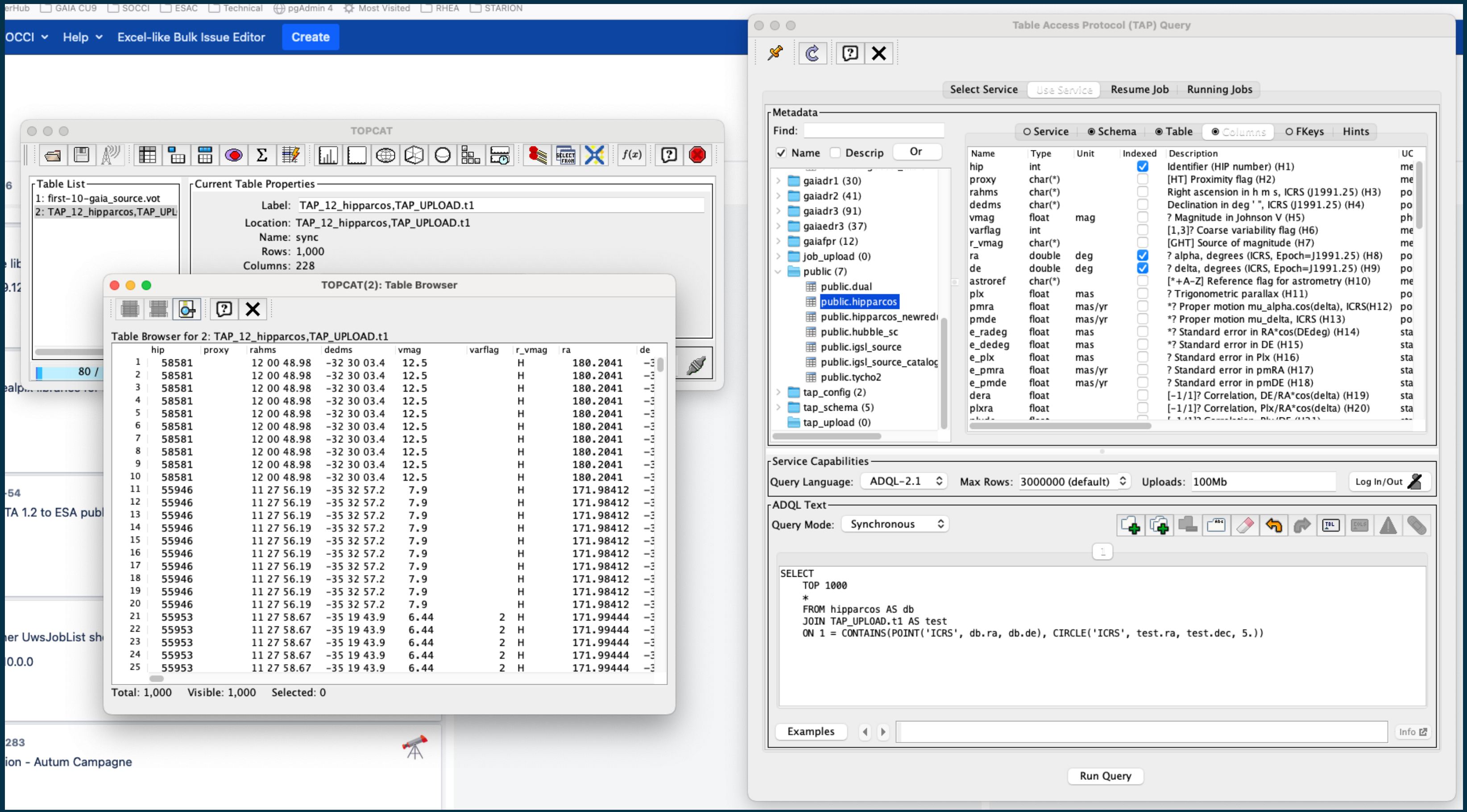


Table List
1: first-10-gaia_source.vot
2: TAP_12_hipparcos,TAP_UPL

Current Table Properties
Label: TAP_12_hipparcos,TAP_UPLOAD.t1
Location: TAP_12_hipparcos,TAP_UPLOAD.t1
Name: sync
Rows: 1,000
Columns: 228

TOPCAT(2): Table Browser

Table Browser for 2: TAP_12_hipparcos,TAP_UPLOAD.t1

	hip	proxy	rahms	dedms	vmag	varflag	r_vmag	ra	de
1	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
2	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
3	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
4	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
5	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
6	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
7	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
8	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
9	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
10	58581		12 00 48.98	-32 30 03.4	12.5		H	180.2041	-3
11	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
12	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
13	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
14	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
15	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
16	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
17	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
18	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
19	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
20	55946		11 27 56.19	-35 32 57.2	7.9		H	171.98412	-3
21	55953		11 27 58.67	-35 19 43.9	6.44	2	H	171.99444	-3
22	55953		11 27 58.67	-35 19 43.9	6.44	2	H	171.99444	-3
23	55953		11 27 58.67	-35 19 43.9	6.44	2	H	171.99444	-3
24	55953		11 27 58.67	-35 19 43.9	6.44	2	H	171.99444	-3
25	55953		11 27 58.67	-35 19 43.9	6.44	2	H	171.99444	-3

Total: 1,000 Visible: 1,000 Selected: 0

Metadata
Find:
✓ Name ☐ Descrip Or

- gaiadr1 (30)
- gaiadr2 (41)
- gaiadr3 (91)
- gaiaedr3 (37)
- gaiafpr (12)
- job_upload (0)
- public (7)
  - public.dual
  - public.hipparcos
  - public.hipparcos_newredı
  - public.hubble_sc
  - public.igsl_source
  - public.igsl_source_catalog
  - public.tycho2
- tap_config (2)
- tap_schema (5)
- tap_upload (0)

Name	Type	Unit	Indexed	Description	UC
hip	int		✓	Identifier (HIP number) (H1)	me
proxy	char(*)			[HT] Proximity flag (H2)	me
rahms	char(*)			Right ascension in h m s, ICRS (J1991.25) (H3)	po
dedms	char(*)			Declination in deg ' ", ICRS (J1991.25) (H4)	po
vmag	float	mag		? Magnitude in Johnson V (H5)	ph
varflag	int			[1,3]? Coarse variability flag (H6)	me
r_vmag	char(*)			[GHT] Source of magnitude (H7)	me
ra	double	deg	✓	? alpha, degrees (ICRS, Epoch=J1991.25) (H8)	po
de	double	deg	✓	? delta, degrees (ICRS, Epoch=J1991.25) (H9)	po
astroref	char(*)			[*+A-Z] Reference flag for astrometry (H10)	me
plx	float	mas		? Trigonometric parallax (H11)	po
pmra	float	mas/yr		*? Proper motion mu_alpha.cos(delta), ICRS(H12)	po
pmde	float	mas/yr		*? Proper motion mu_delta, ICRS (H13)	po
e_radeg	float	mas		*? Standard error in RA*cos(DEdeg) (H14)	sta
e_dedeg	float	mas		*? Standard error in DE (H15)	sta
e_plx	float	mas		? Standard error in Plx (H16)	sta
e_pmra	float	mas/yr		? Standard error in pmRA (H17)	sta
e_pmde	float	mas/yr		? Standard error in pmDE (H18)	sta
dera	float			[-1/1]? Correlation, DE/RA*cos(delta) (H19)	sta
plxra	float			[-1/1]? Correlation, Plx/RA*cos(delta) (H20)	sta

Service Capabilities
Query Language: ADQL-2.1 Max Rows: 3000000 (default) Uploads: 100Mb Log In/Out

ADQL Text
Query Mode: Synchronous

```
SELECT
    TOP 1000
    *
    FROM hipparcos AS db
    JOIN TAP_UPLOAD.t1 AS test
    ON 1 = CONTAINS(POINT('ICRS', db.ra, db.de), CIRCLE('ICRS', test.ra, test.dec, 5.))
```

”On-the-fly" Upload Capabilities

Using a job result in a query (I)

- Typical use case: Upload local data to an external TAP
- This mechanism allows to run a query against a job executed in our archive, using the unique alpha-numeric code assigned to each job

```
SELECT upload.*, catwise.*
FROM tap_upload.job16415687691150 AS upload
JOIN "II/365/catwise" AS catwise ON 1 = CONTAINS(
                    POINT('ICRS', catwise.RA_ICRS, catwise.DE_ICRS),
                    CIRCLE('ICRS', upload.ra, upload.dec, 1. / 3600.))
```

”On-the-fly" Upload Capabilities

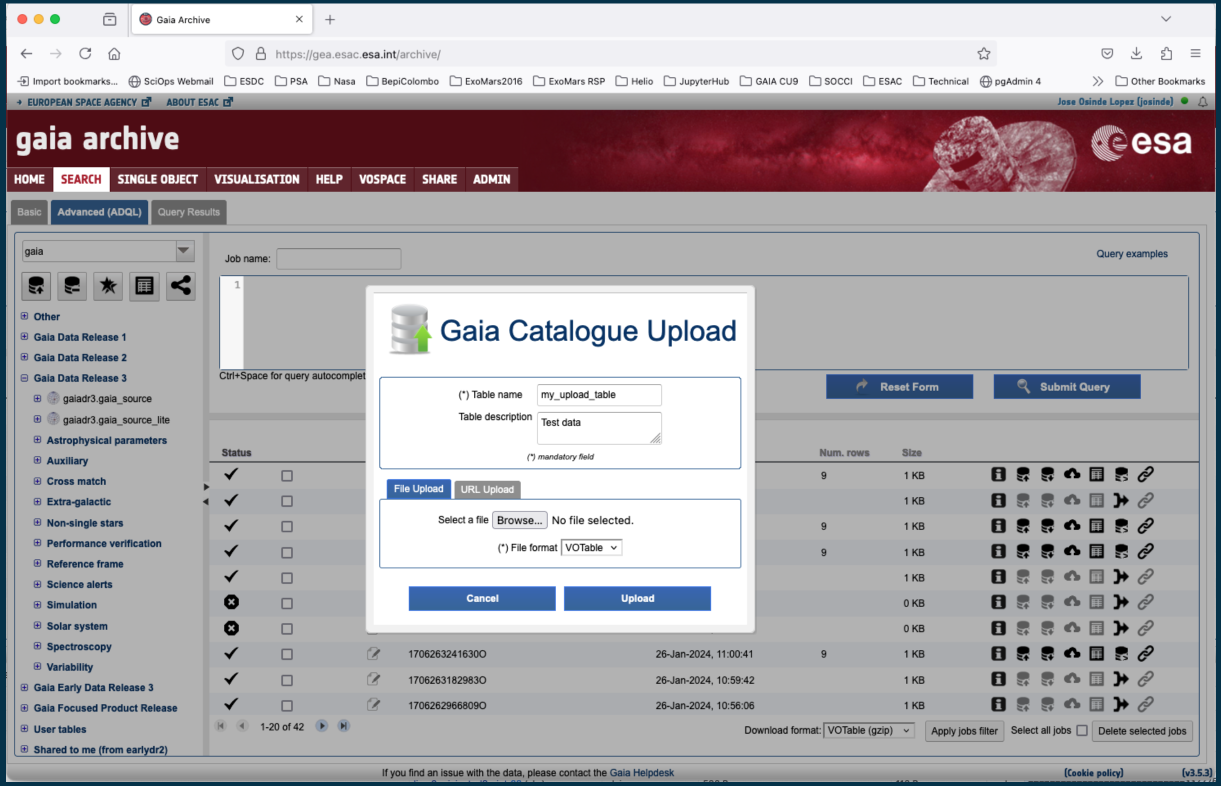
Using a job results into a query (II)

- No external TAP is involved in this case
- Data is uploaded to the database as a new table persistent during the execution of the query
- Use a different schema to store data into the database: JOB_UPLOAD
- Simple example:

```
SELECT * FROM job_upload.”job1641568769115O”
```

Persistent uploads

- Service accessible only to registered users
- Users can upload data from local files or URLs
- User can also upload job results stored in the archive
- Data stored in a schema associated to the user
- Once uploaded it can be used as any other table in the archive and shared with other users if required.



Gaia Catalogue Upload

(*) Table name: my_upload_table

Table description: Test data

(*) mandatory field

File Upload | URL Upload

Select a file: Browse... No file selected.

(*) File format: VOTable

Cancel | Upload

Status	Job	Date	Num. rows	Size
✓			9	1 KB
✓				1 KB
✓			9	1 KB
✓			9	1 KB
✓				1 KB
✖				0 KB
✖				0 KB
✓	1706263241630O	26-Jan-2024, 11:00:41	9	1 KB
✓	1706263182983O	26-Jan-2024, 10:59:42		1 KB
✓	1706262966809O	26-Jan-2024, 10:56:06		1 KB

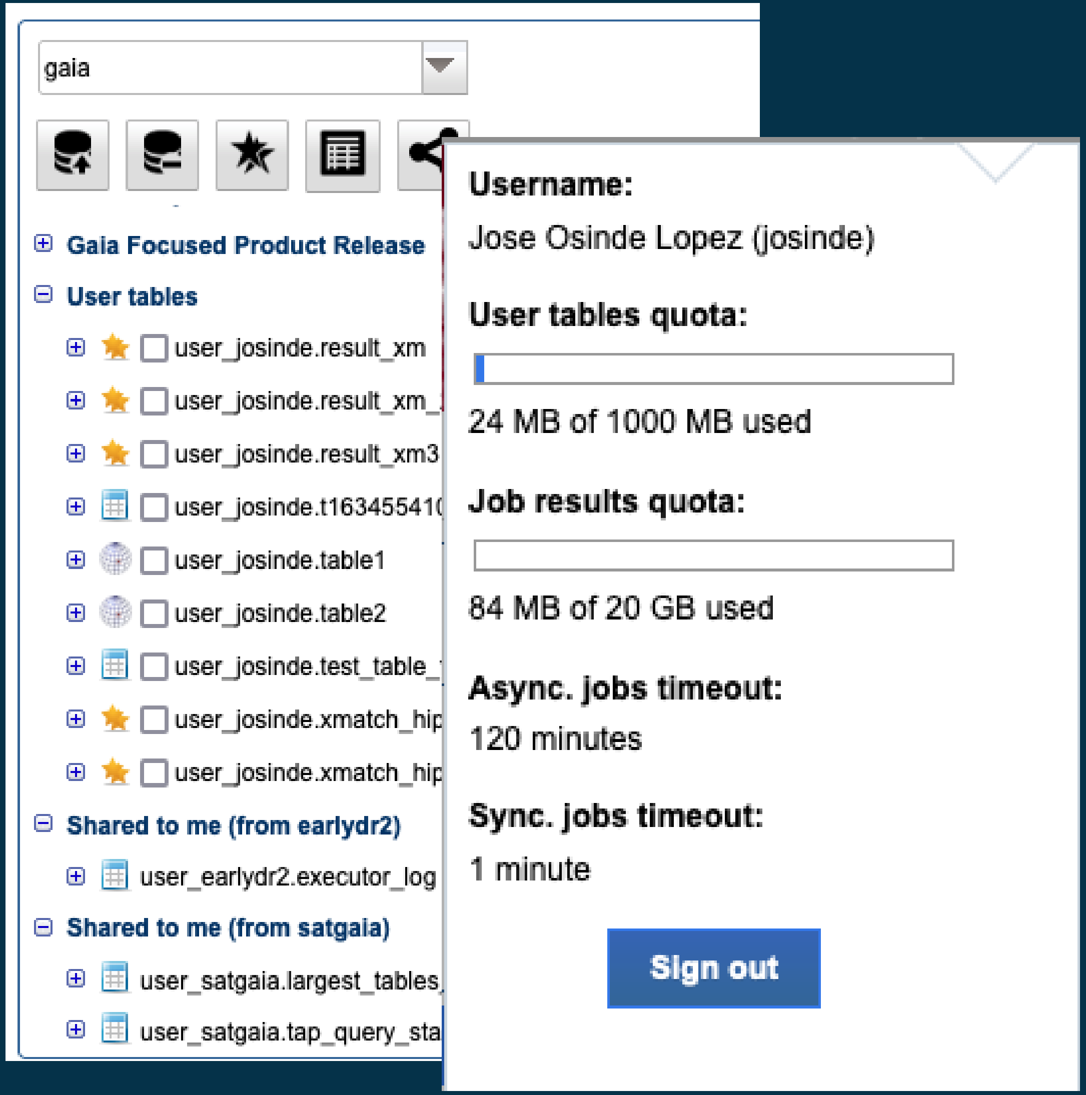
1-20 of 42 — Download format: VOTable (gzip) — Apply jobs filter — Select all jobs — Delete selected jobs

If you find an issue with the data, please contact the Gaia Helpdesk (Cookie policy) (v3.5.3)

Persistent uploads

Private areas involves:

- User authentication
    - Integration with directory services (LDAP)
- Quota management
- Space organization in the database (schemas)



gaia

Gaia Focused Product Release

User tables

- user_josinde.result_xm
- user_josinde.result_xm_
- user_josinde.result_xm3
- user_josinde.t163455410
- user_josinde.table1
- user_josinde.table2
- user_josinde.test_table_
- user_josinde.xmatch_hip
- user_josinde.xmatch_hip

Shared to me (from earlydr2)

- user_earlydr2.executor_log

Shared to me (from satgaia)

- user_satgaia.largest_tables
- user_satgaia.tap_query_sta

Username: Jose Osinde Lopez (josinde)

User tables quota: 24 MB of 1000 MB used

Job results quota: 84 MB of 20 GB used

Async. jobs timeout: 120 minutes

Sync. jobs timeout: 1 minute

Sign out

Upload endpoints

Command line examples:

- Upload a table from a file:

```
curl -k -b cookies.txt -X POST -F FILE=@file.name
                               -F TABLE_NAME=table_name
                                       https://host:port/context/Upload
```

- Delete a user table from the database:

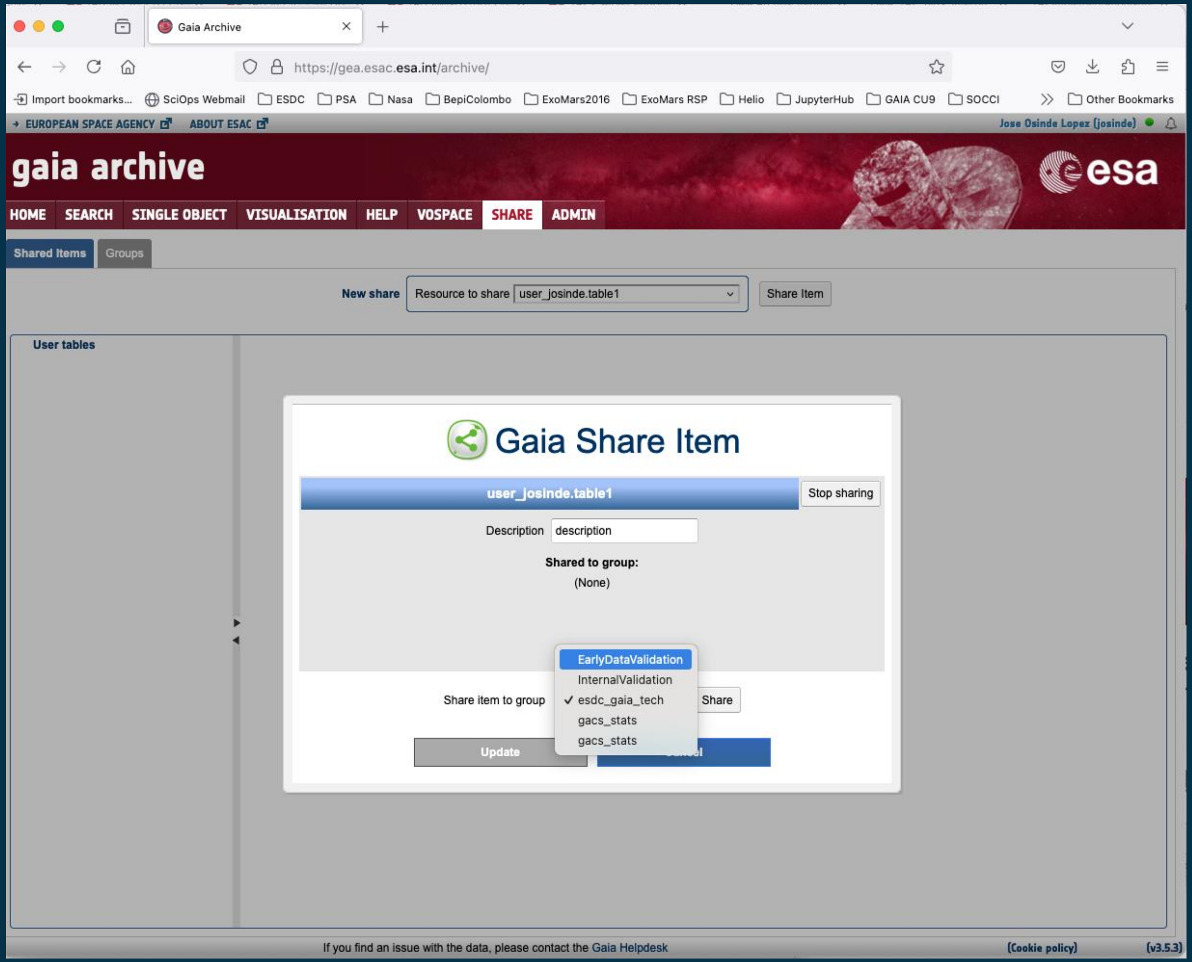
```
curl -k -b cookies.txt -X POST -F TABLE_NAME=table_name
                               -F DELETE=TRUE
                               -F FORCE_REMOVAL=TRUE
                                       https://host:port/context/Upload
```

Persistent uploads

Sharing data involves:

- Implement management tools:
  - Create/update/delete user or groups
  - Handle the reference between these two elements
  - Register resources and link these resources with one or more groups

- New end-points supporting this functionality



New share — Resource to share: user_josinde.table1

User tables

Gaia Share Item

user_josinde.table1

Description: description

Shared to group:
(None)

Share item to group: ✓ esdc_gaia_tech

If you find an issue with the data, please contact the Gaia Helpdesk

Gaia Archive
https://gea.esac.esa.int/archive/
Import bookmarks... SciOps Webmail ESDC PSA Nasa BepiColombo ExoMars2016 ExoMars RSP Helio JupyterHub GAIA CU9 SOCCI Other Bookmarks
EUROPEAN SPACE AGENCY ABOUT ESAC
Jose Osinde Lopez (josinde)
gaia archive
esa
HOME SEARCH SINGLE OBJECT VISUALISATION HELP VOSPACE SHARE ADMIN
Shared Items
Groups
Share Item
Stop sharing
EarlyDataValidation
InternalValidation
gacs_stats
gacs_stats
Share
Update
Cancel
(Cookie policy)
(v3.5.3)

Share endpoint

Create group:

```
http://host:port/context/share?action=CreateOrUpdateGroup&title=My+group&users_list=user1,user2
```

Add user to a group:

```
http://host:port/context/share?action=CreateUserGroup&group_id=group_id&user_id=user_id
```

Create a shared item:

```
http://host:port/context/share?action=CreateOrUpdateItem&resource_type=table
      &title=My+table&description=Description&items_list=group_id_1,Group,Read|user_id_1,User,Write
```

Create a shared item relation:

```
http://host:port/context/share?action=CreateItemRelation&resource_id=resource_id
       &resource_type=0&share_to_id=identifier&share_type=Group&share_mode=Read
```

Questions / feedback

Thank you for your attention