GENIUS

The Gaia archive

gaia

Gaia DPAC
Data Processing & Analysis Consortium

X. Luri
Universitat de Barcelona

Archive current contents

External Catalogues

• Hipparcos & Hipparcos new red.	1.2×10^6	rows
• IGSL (Initial Gaia Source List)	1.2×10^9	rows
• 2MASS	4.7×10^8	rows
• Tycho2	2.5×10^6	rows
• UCAC4	1.1×10^8	rows
• Hubble Source Catalogue v1.0	2.9×10^7	rows

Gaia

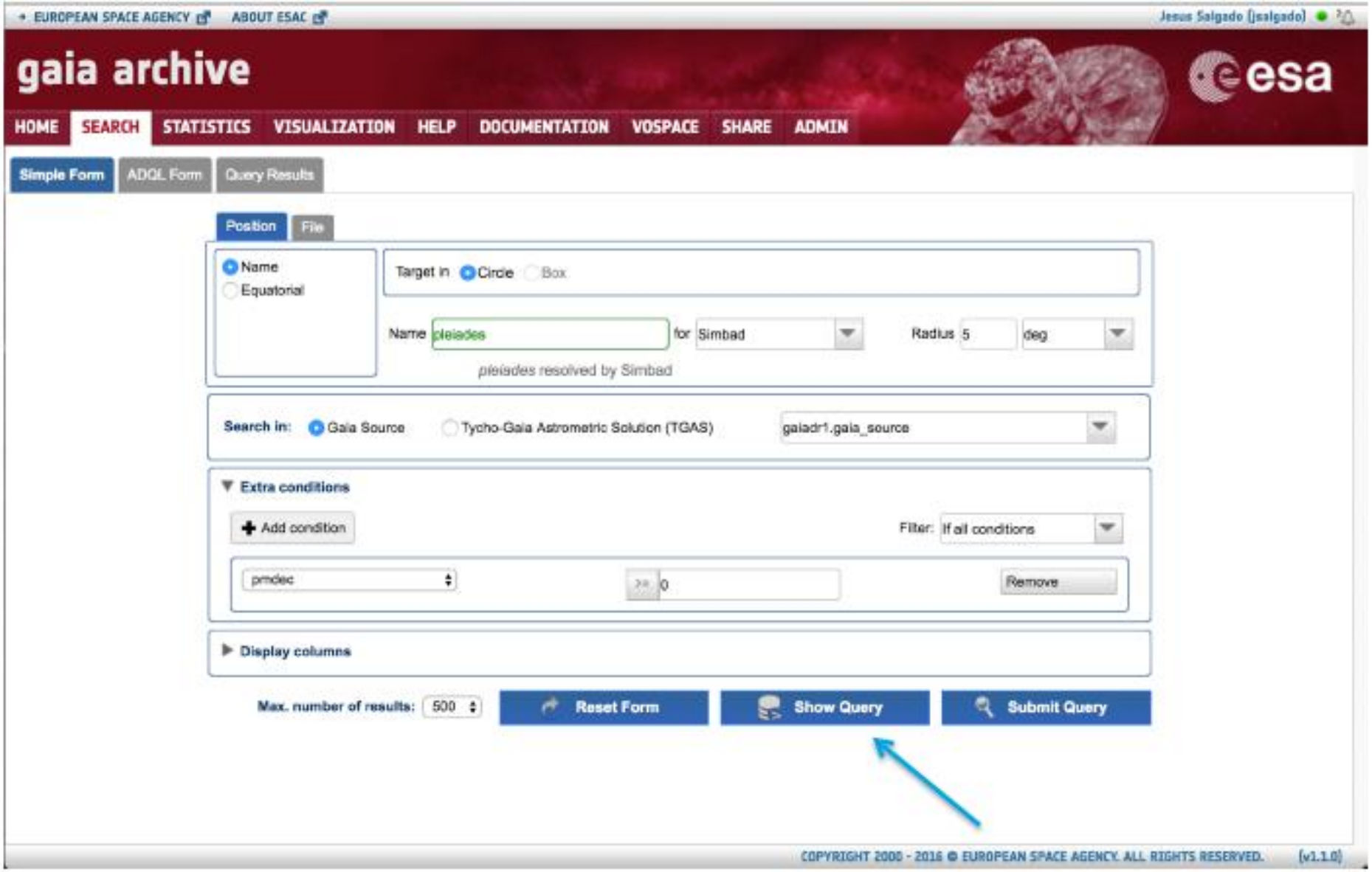
• Gaia DR1 catalogue	1.1×10^9	rows
• TGAS	2.0×10^6	rows

Crossmatches

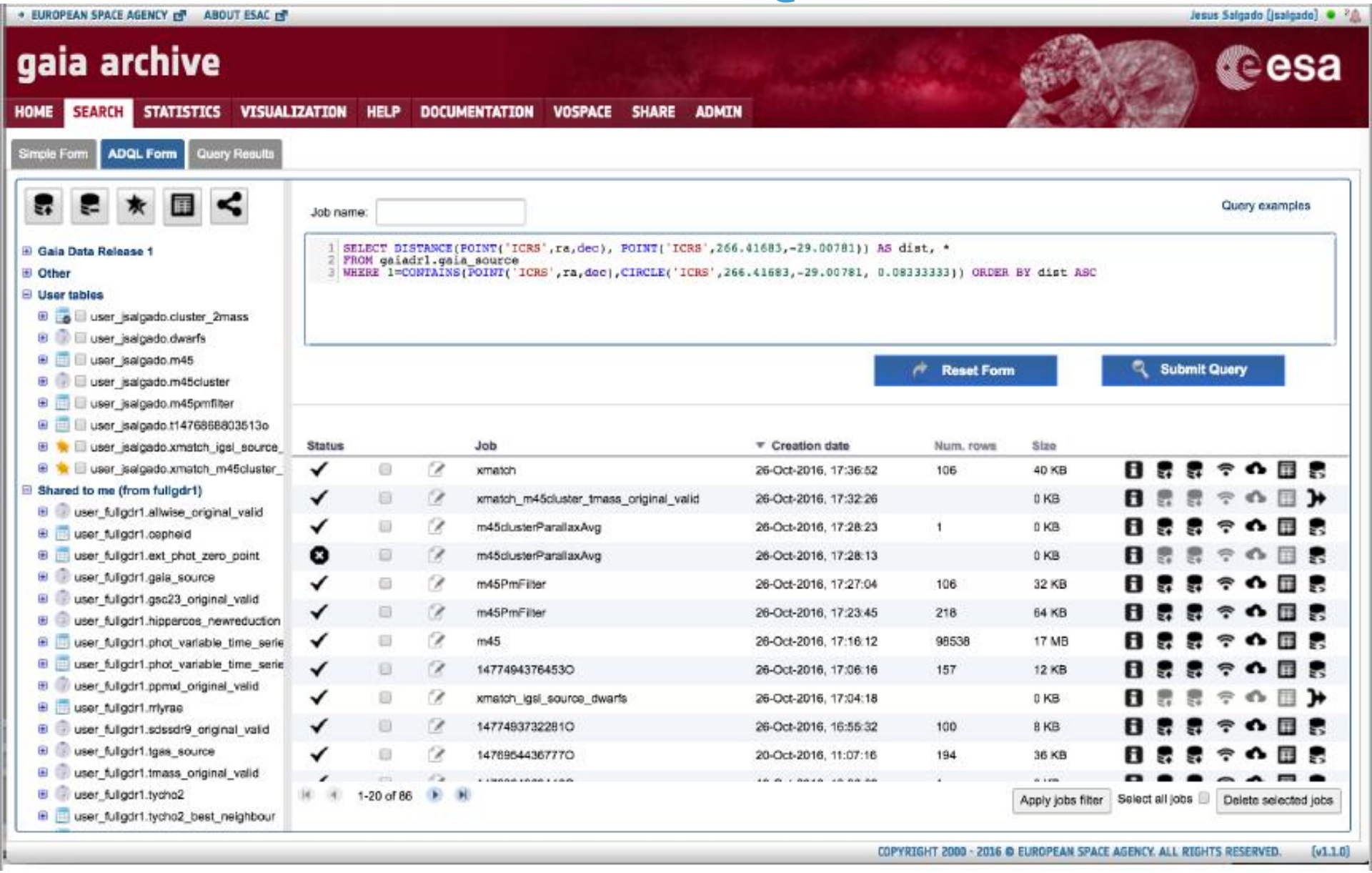
- Crossmatch tables between Hipparcos, 2MASS, Tycho2... and Gaia expressed as neighbourhood and best neighbour, e.g:

• AllWise-Gaia neighbourhood	3.1×10^8	rows

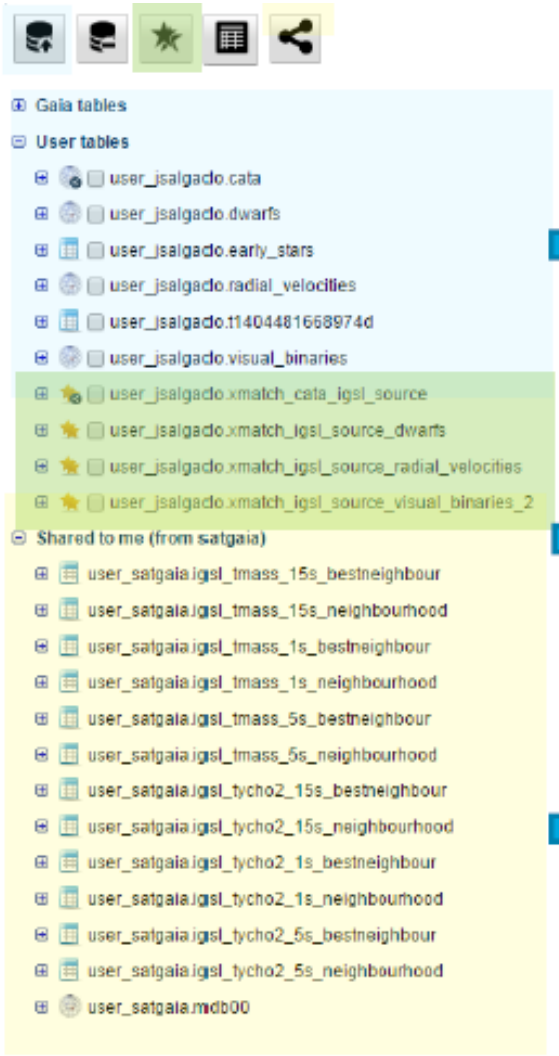
Archive simple form

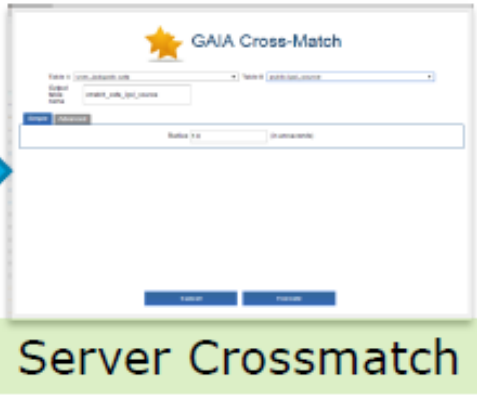


Archive ADQL form



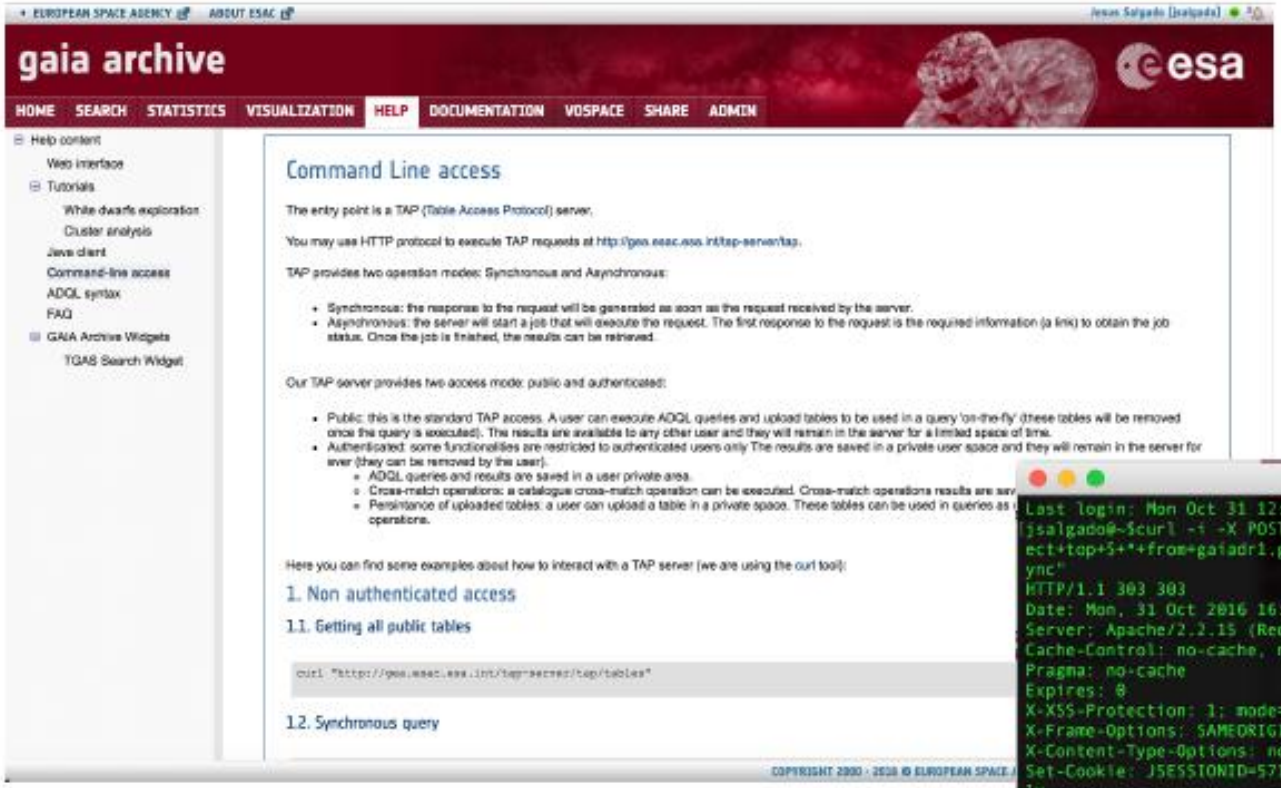
Archive user schemas



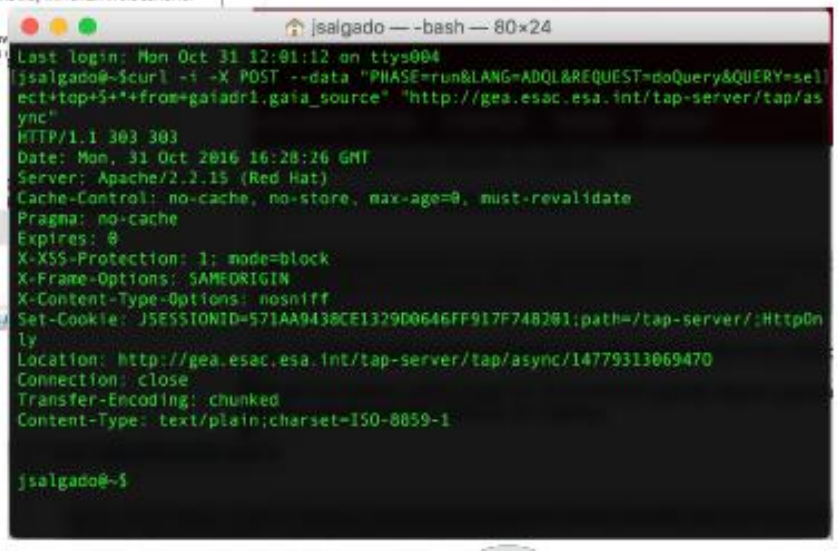


ESA UNCLASSIFIED - Releasable to the Public

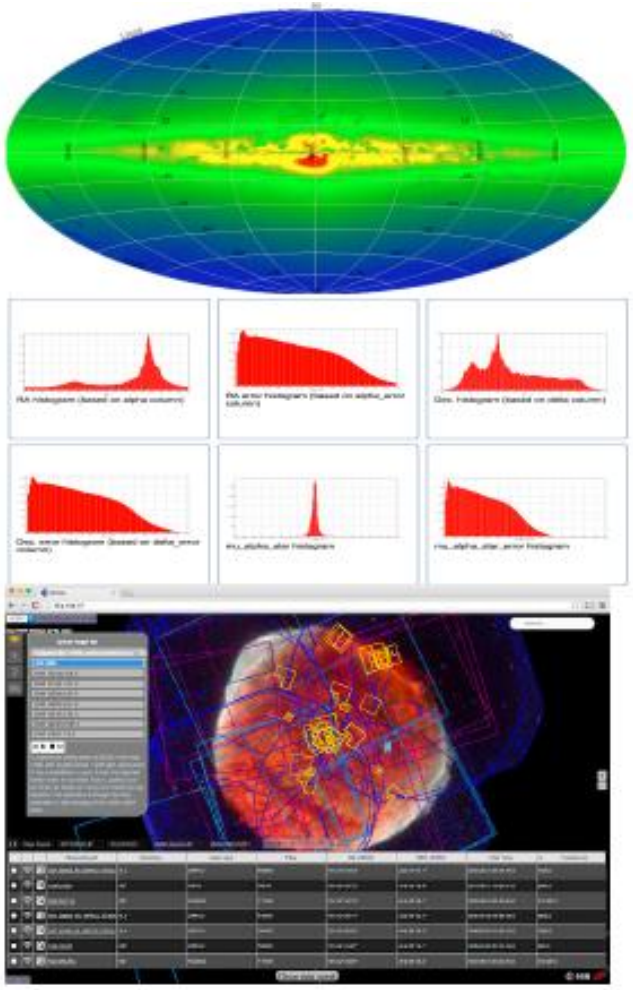
Archive CLI


EUROPEAN SPACE AGENCY
ABOUT ESAC
gaia archive
esa
HOME
SEARCH
STATISTICS
VISUALIZATION
HELP
DOCUMENTATION
VOSPACE
SHARE
ADMIN
Help content
Web interface
Tutorials
White dwarfs exploration
Cluster analysis
Java client
Command-line access
ADQL syntax
FAQ
GAIA Archive Widgets
TGAS Search Widget
Command Line access
The entry point is a TAP (Table Access Protocol) server.
You may use HTTP protocol to execute TAP requests at http://gea.esac.esa.int/tap-server/tap.
TAP provides two operation modes: Synchronous and Asynchronous:
Synchronous: the response to the request will be generated as soon as the request received by the server.
Asynchronous: the server will start a job that will execute the request. The first response to the request is the required information (a link) to obtain the job status. Once the job is finished, the results can be retrieved.
Our TAP server provides two access mode: public and authenticated:
Public: this is the standard TAP access. A user can execute ADQL queries and upload tables to be used in a query 'on-the-fly' (these tables will be removed once the query is executed). The results are available to any other user and they will remain in the server for a limited space of time.
Authenticated: some functionalities are restricted to authenticated users only. The results are saved in a private user space and they will remain in the server for ever (they can be removed by the user).
ADQL queries and results are saved in a user private area.
Cross-match operations: a catalogue cross-match operation can be executed. Cross-match operations results are sav
Persistance of uploaded tables: a user can upload a table in a private space. These tables can be used in queries as
operations.
Here you can find some examples about how to interact with a TAP server (we are using the curl tool):
1. Non authenticated access
1.1. Getting all public tables
curl "http://gea.esac.esa.int/tap-server/tap/tables"
1.2. Synchronous query
COPYRIGHT 2000 - 2016 © EUROPEAN SPACE

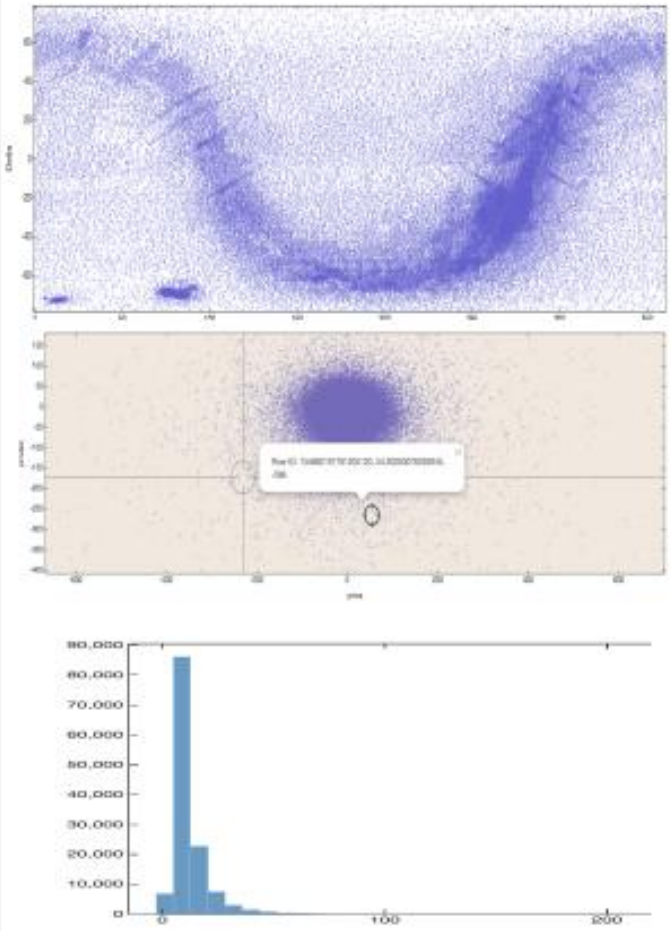
PY
Java

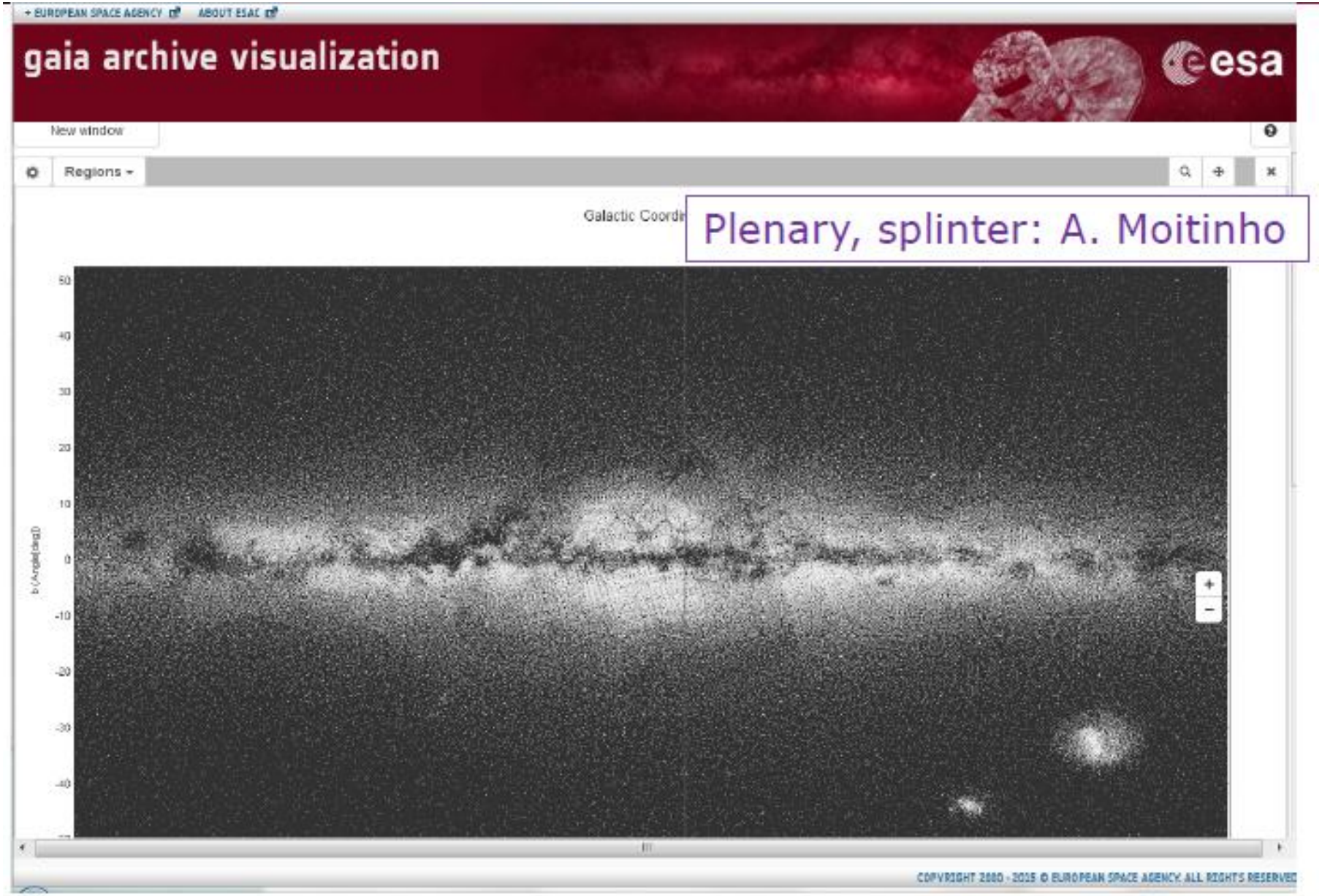

jsalgado — -bash — 80×24
Last login: Mon Oct 31 12:01:12 on ttys004
[jsalgado@~$curl -i -X POST --data "PHASE=run&LANG=ADQL&REQUEST=doQuery&QUERY=select+top+5+*+from+gaiadr1.gaia_source" "http://gea.esac.esa.int/tap-server/tap/async"
HTTP/1.1 303 303
Date: Mon, 31 Oct 2016 16:28:26 GMT
Server: Apache/2.2.15 (Red Hat)
Cache-Control: no-cache, no-store, max-age=0, must-revalidate
Pragma: no-cache
Expires: 0
X-XSS-Protection: 1; mode=block
X-Frame-Options: SAMEORIGIN
X-Content-Type-Options: nosniff
Set-Cookie: JSESSIONID=571AA9438CE1329D0646FF917F748201;path=/tap-server/;HttpOnly
Location: http://gea.esac.esa.int/tap-server/tap/async/1477931306947O
Connection: close
Transfer-Encoding: chunked
Content-Type: text/plain;charset=ISO-8859-1
jsalgado@~$

Archive visualization



- Visualization is a need
- Statistics provide holistic views. Big data techniques
- Validated static stats
- On-the-fly visualization (Lisbon University)
- ESASky




EUROPEAN SPACE AGENCY
ABOUT ESAC
gaia archive visualization
esa
New window
Regions
Plenary, splinter: A. Moitinho
COPYRIGHT 2000 - 2015 © EUROPEAN SPACE AGENCY. ALL RIGHTS RESERVED

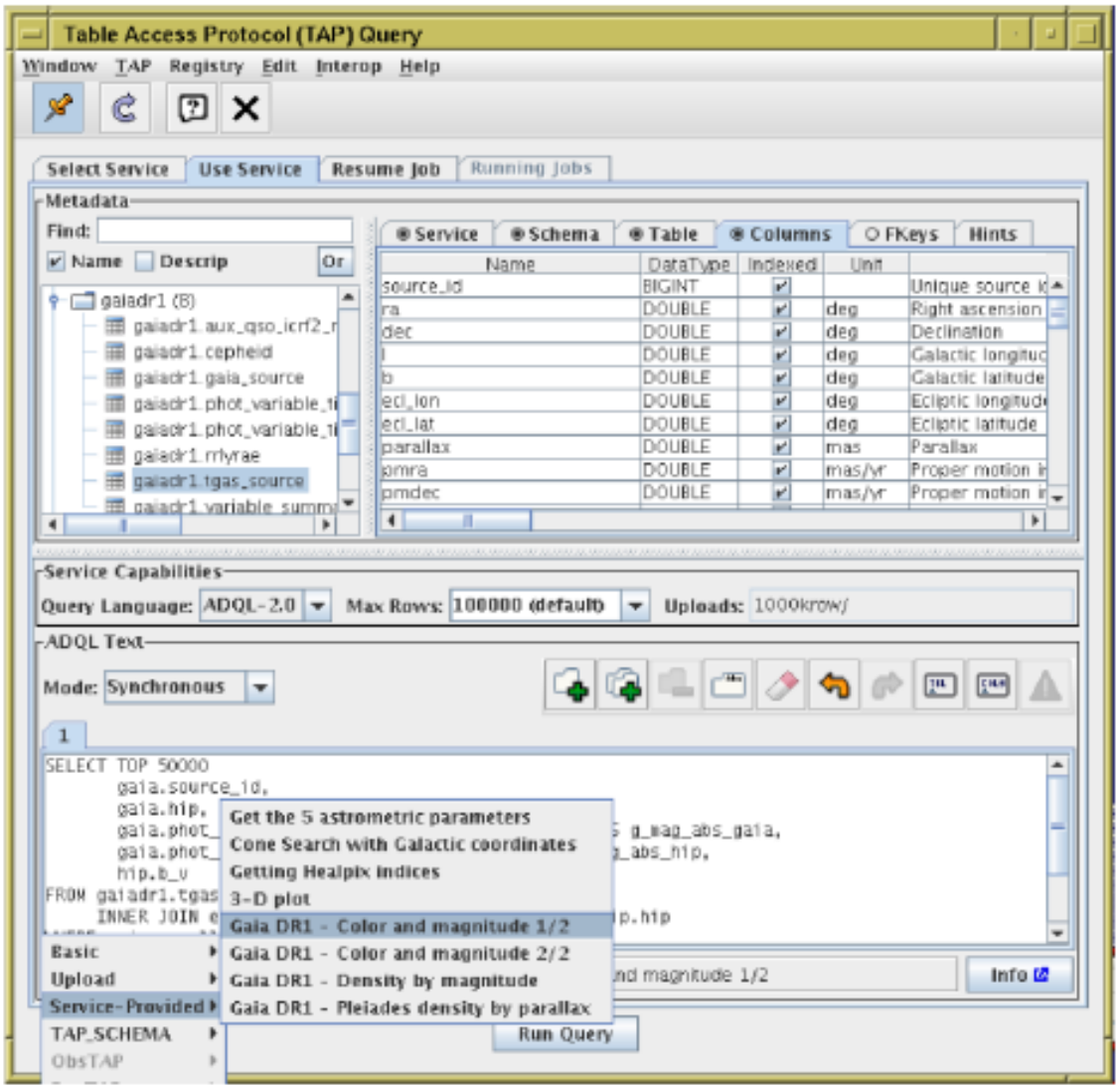


Table Access Protocol (TAP) Query
Window TAP Registry Edit Interop Help
Select Service
Use Service
Resume Job
Running Jobs
Metadata
Find:
Name
Descrip
Service
Schema
Table
Columns
FKeys
Hints
gaiadr1 (8)
gaiadr1.aux_qso_icrf2_r
gaiadr1.cepheid
gaiadr1.gaia_source
gaiadr1.phot_variable_1
gaiadr1.phot_variable_1
gaiadr1.rrlyrae
gaiadr1.tgas_source
gaiadr1.variable_summ
Name
DataType
Indexed
Unit
source_id
BIGINT
Unique source i
ra
DOUBLE
deg
Right ascension
dec
DOUBLE
deg
Declination
l
DOUBLE
deg
Galactic longitud
b
DOUBLE
deg
Galactic latitude
ecl_lon
DOUBLE
deg
Ecliptic longitud
ecl_lat
DOUBLE
deg
Ecliptic latitude
parallax
DOUBLE
mas
Parallax
pmra
DOUBLE
mas/yr
Proper motion i
pmdec
DOUBLE
mas/yr
Proper motion i
Service Capabilities
Query Language: ADQL-2.0
Max Rows: 100000 (default)
Uploads: 1000krow/
ADQL Text
Mode: Synchronous
1
SELECT TOP 50000
gaia.source_id,
gaia.hip,
gaia.phot_
gaia.phot_
hip.b_v
FROM gaiadr1.tgas
INNER JOIN e
g_mag_abs_gaia,
_abs_hip,
p.hip
Get the 5 astrometric parameters
Cone Search with Galactic coordinates
Getting Healpix indices
3-D plot
Gaia DR1 - Color and magnitude 1/2
Gaia DR1 - Color and magnitude 2/2
Gaia DR1 - Density by magnitude
Gaia DR1 - Pleiades density by parallax
Basic
Upload
Service-Provided
TAP_SCHEMA
ObsTAP
nd magnitude 1/2
Info
Run Query
Plane Plot
Window Layers Subsets Plot Export Help
g_mag_abs_gaia
b_v / mag
Frame
Legend
Axes
Position
Subsets
Form
Table: 19: TAP_10_gaiadr1.tgas_source,extcat.hipparcos
X: b_v
Y: g_mag_abs_gaia
Position: 1.67, 1.9
Count: 43,517 / 43,519
Select
Pan X/Y
Stretch X/Y
Frame X/Y
Zoom

Archive documentation

- Gaia Archive. Includes help and tutorials
  - https://archives.esac.esa.int/gaia
- Gaia DR1 papers http://www.cosmos.esa.int/web/gaia/dr1#A&A
- Online documentation (361 pages)
  - http://gaia.esac.esa.int/documentation/GDR1/index.html
- Data model documentation
  - https://gaia.esac.esa.int/documentation/GDR1/datamodel/
- ADQL: GAVO short course, UK ROE cookbook
  - http://docs.g-vo.org/adql-gaia/html/index.html
  - https://gaia.ac.uk/science/gaia-data-release-1/adql-cookbook
- Gaia Helpdesk. https://support.cosmos.esa.int/gaia/

Future: data mining

AVI

- **Added Value Interface**
  - Provides web-interface
  - Executes pipeline
- Hosted in **AVI Container**
- Runs AVI framework

Portal Systems

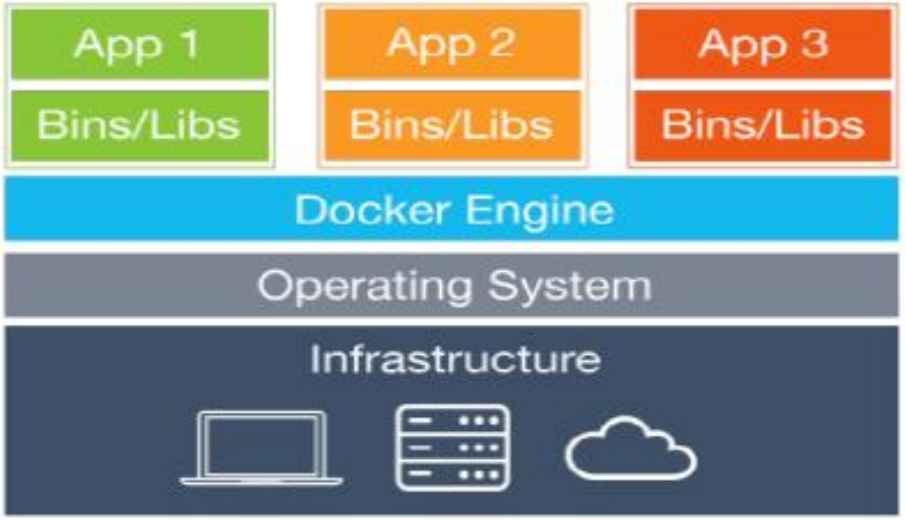
- **Public point of access**
  - Web/REST
- Manages:
  - Users/Roles
  - Authentication
  - AVI deployment
  - Job scheduling

AVI Infrastructure

- Hosts AVI containers

Supporting Systems

- Monitors GAVIP
- Hosts message bus
  - Several queues
  - Distributed



1. Improve "code to the data" paradigm
2. GAVI portal
3. Data mining (Spark? Hadoop?)