

The 25 Years of TNG through the Archival Data

Andrea Bignamini and IA2 Team
INAF

Telescopio Nazionale Galileo: 25 years of Astronomy in La Palma
19 October 2021



IA2 Main Goals



Centro Italiano Archivi Astronomici (IA2) main goals are :

- archiving systems
- safety
- data curation and preservation
- distribution over geographical distributed sites
- providing services and tools (collaboration tools, workflow, etc..)
- data publication in the VO

IA2 manages data of several projects:



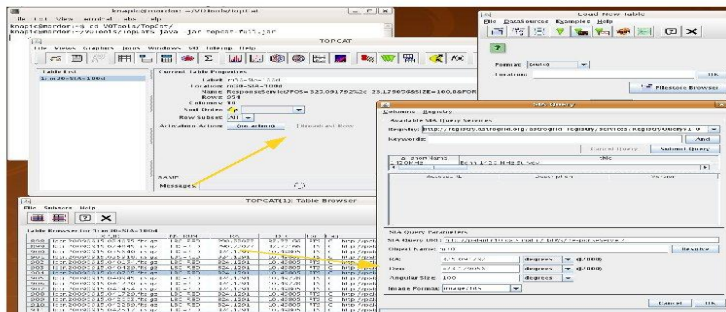
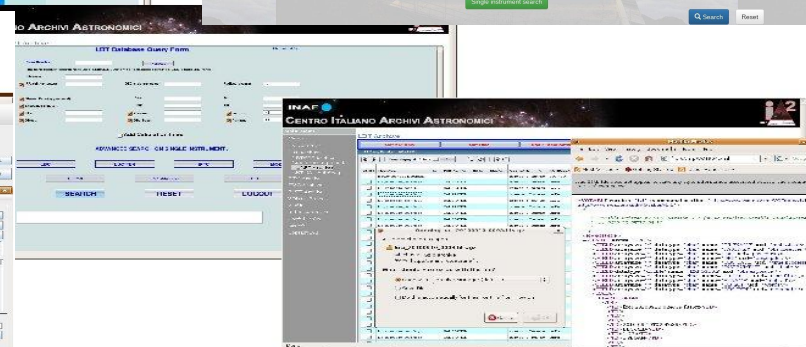
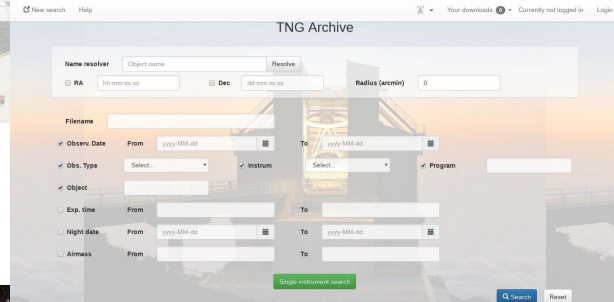
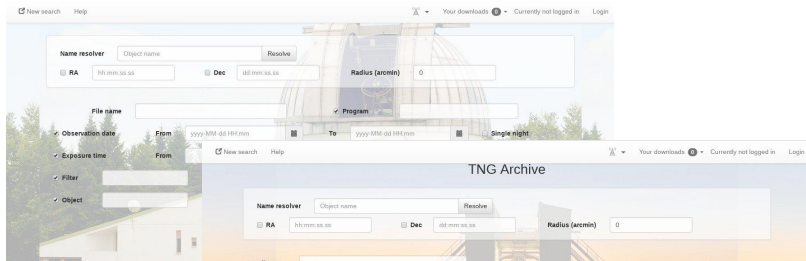
- Telescopes: TNG, LBT, Asiago, etc.
- Surveys: raw and/or calibrated
- Exoplanets (GAPS)
- Simulations (ITVO)
- EU Projects (GENIUS, VIALACTEA, INDIGO, ASTERICS)
- Radio Archives
- Educational (SVAS)

IA2 is part of INAF-ICT

IA2 Services for the Astronomical Community



VO Service & TopCat

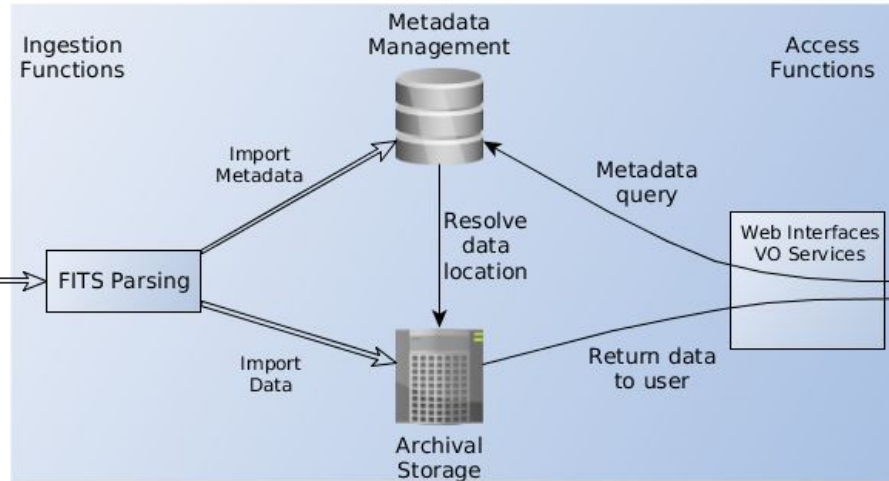


TNG Archive Schema

Telescope
Data Production



Archive
Data Stewardship



Final User
Data Analysis



Brief History of TNG Archive - part 1

- 1993 CRA nominated a task force named “TNG Archives Working Group” to define system requirements and basic architecture for the “Archives at the TNG” and the “General Archive”
- 1995 Work concerning the Archives at the Telescope (AaT) started at the Astronomical Observatory of Trieste (OATs)
- 1999 AaT system officially delivered at CGG

Brief History of TNG Archive - part 1

- 1993 CRA nominated a task force named “TNG Archives Working Group” to define system requirements and basic architecture for the “Archives at the Telescope (AaT) system”
- 1995 Work concerning the Archives at the Telescope (AaT) started at OATs Trieste (OATs)
- 1999 AaT system officially delivered at CGG
- 2001 Pilot study for Long-Term Archive (LoTar) started at OATs
 - Persistent archive
 - Provide access to users and to the astronomical community

A PILOT PROJECT FOR THE IMPLEMENTATION OF THE ITALIAN “TELESCOPIO NAZIONALE GALILEO” LONG-TERM ARCHIVE

Principal Investigator : Fabio Pasián

OATg People:

Riccardo Smareglia, Alessandro Caproni, Andrea Zacchei, Claudio Vuerli

Coordinating Institute :

Osservatorio Astronomico di Trieste;

Participating Institutes :

Osservatorio Astronomico di Roma;
Osservatorio Astronomico di Catania;
Osservatorio Astronomico di Capodimonte;
Centro “Galileo Galilei” (TNG) – La Palma;
Istituto di Radio-Astronomia CNR – Bologna;
Centre Données astronomiques de Strasbourg;
Space Telescope European Coordinating Facility, ESO, Garching



20-12-2001

A Pilot Project for the implementation of
the TNG Long Term Archive


1



Brief History of TNG Archive - part 1

- 1993 CRA nominated a task force named “TNG Archives Working Group” to define system requirements and basic architecture for the “Archives at the TNG” and the “General Archive”
- 1995 Work concerning the Archives at the Telescope (AaT) started at the Astronomical Observatory of Trieste (OATs)
- 1999 AaT system officially delivered at CGG
- 2001 Pilot study for Long-Term Archive (LoTAr) started at OATs
 - Persistent archive
 - Provide access to users and to the astronomical community
- 2003 IA2 created with the main goal to set up a Long-Term Archive for TNG

Brief History of TNG Archive - part 1

	<p>“CENTER FOR ASTRONOMICAL ARCHIVES” <i>Response to the INAF Announcement of Opportunity</i></p>	<p>Document No. Issue/Rev. No. Date Page</p>	<p>IA²-OATs-RF-001 1.0 10 November 2003 29 of 87</p>
---	--	--	---

2.2.2.3 Ingestion of TNG data

2.2.2.3.1 Ingestion of TNG scientific data

Files containing scientific exposures are sent by CGG to LTA via mail or courier [22]. The ingestion procedure reads all CD-Rs loaded on the juke-box and dumps their content on magnetic disk to allow on-line access.

Brief History of TNG Archive - part 1

About 1000 CD-ROM



The AWS data transport solution

AWS Snowball—petabyte scale data transport



E-ink shipping
label

RE:INVENT 2015 LAUNCH

- Rugged 8.5G impact case
- Rain and dust resistant
- Data encryption end-to-end
- 50 TB capacity/10G network
- E-ink label

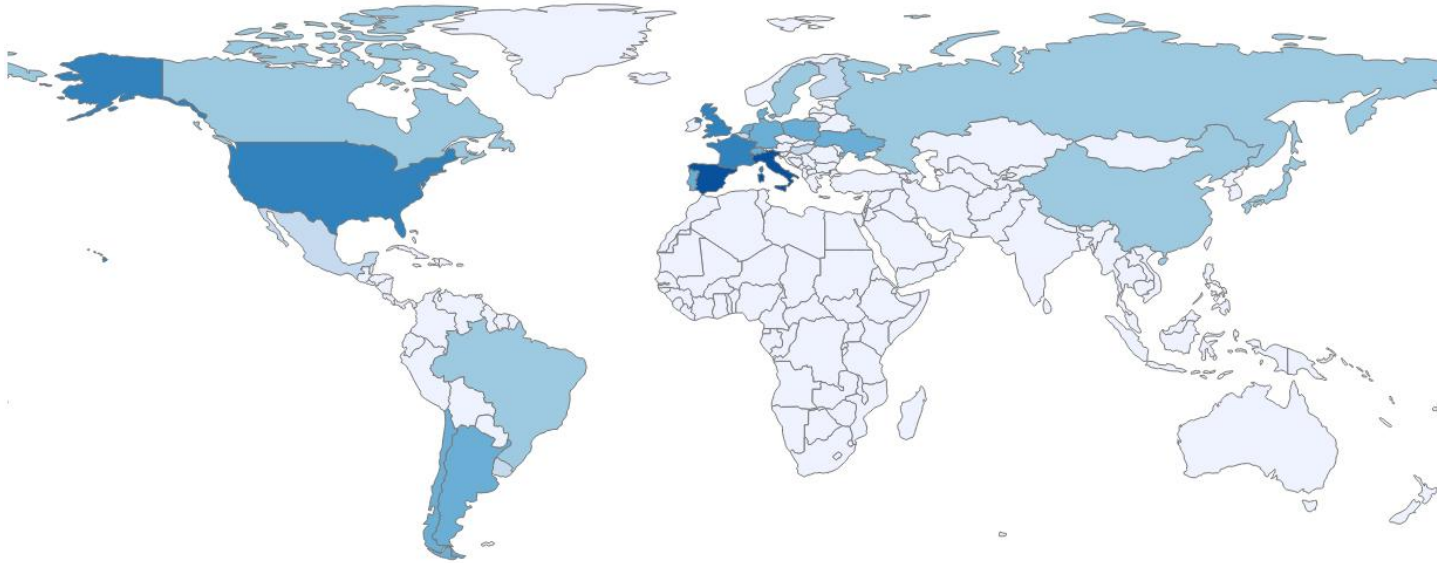


Credit to CM Vahid
Principal Solutions
Architect AWS AI

Brief History of TNG Archive - part 2

- 2006 TNG LTA started at IA2
 - Based on DBMS Oracle 8i Enterprise Edition
 - Already with interoperability features (e.g. SIMBAD Name Resolver)
- 2012 Second phase of TNG Archive
 - New Java Data Ingester + MySQL + new web interfaces
- 2015 NADIR (New Archiving Distributed InfrastructuRe) and Third phase of TNG Archive
 - Archiving functions treated as “devices” managed by Tango Controls
 - Modular, scalable and designed to manage large distributed archives, multiversion of observations, and data distribution according to well defined access policy
 - On the top of the archive many other services have been developed:
 - Apogeo and Portal Generator
 - User Management Tools: Archive Program Manager (APM), Group Membership Service (GMS), Single Sign-On (SSO)

TNG Archive Map of Accesses



ITA: 2879
ESP: 1367
GBR: 517
USA: 488
FRA: 288

NLD: 161
CHE: 148
ARG: 94
DEU: 92
CHL: 76

DNK: 71
PRT: 70
POL: 68
UKR: 61
BRA: 28

RUS: 26
SWE: 24
CHN: 24
CAN: 17
JPN: 14

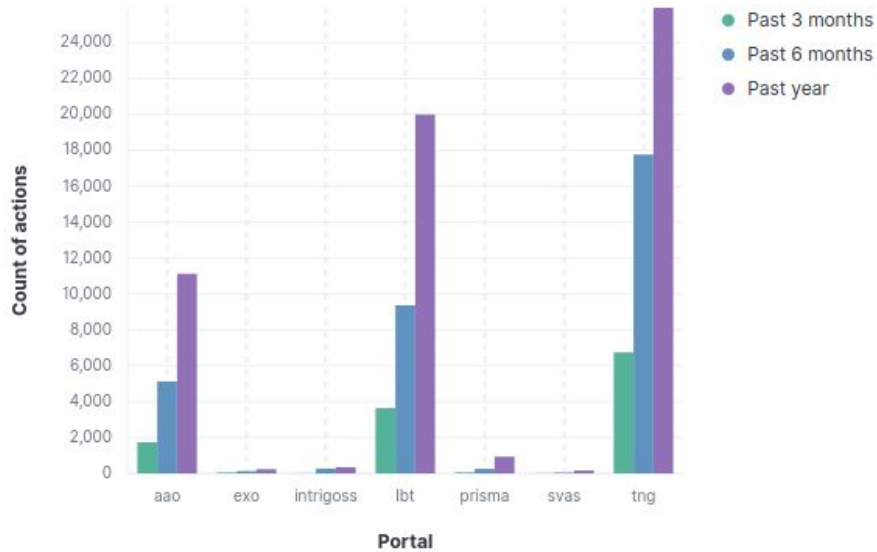
URY: 8
MEX: 6
BEL: 6
ISR: 5
HUN: 5

FIN: 3
ROU: 2
AUS: 2
TWN: 1
MUS: 1

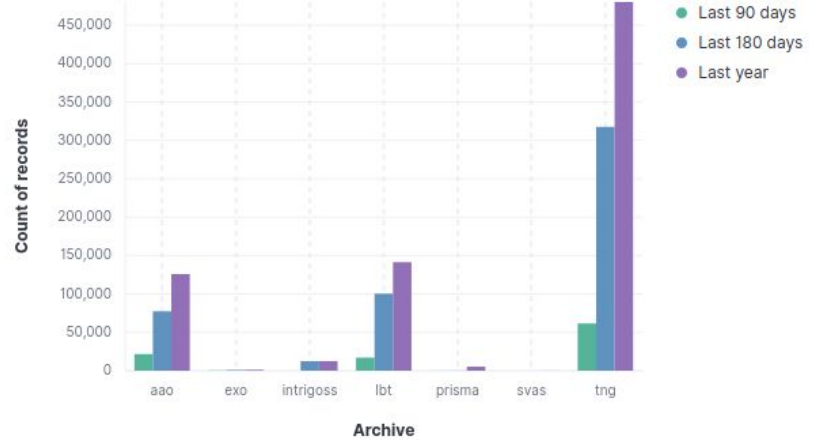
LTU: 1
IND: 1
ARM: 1

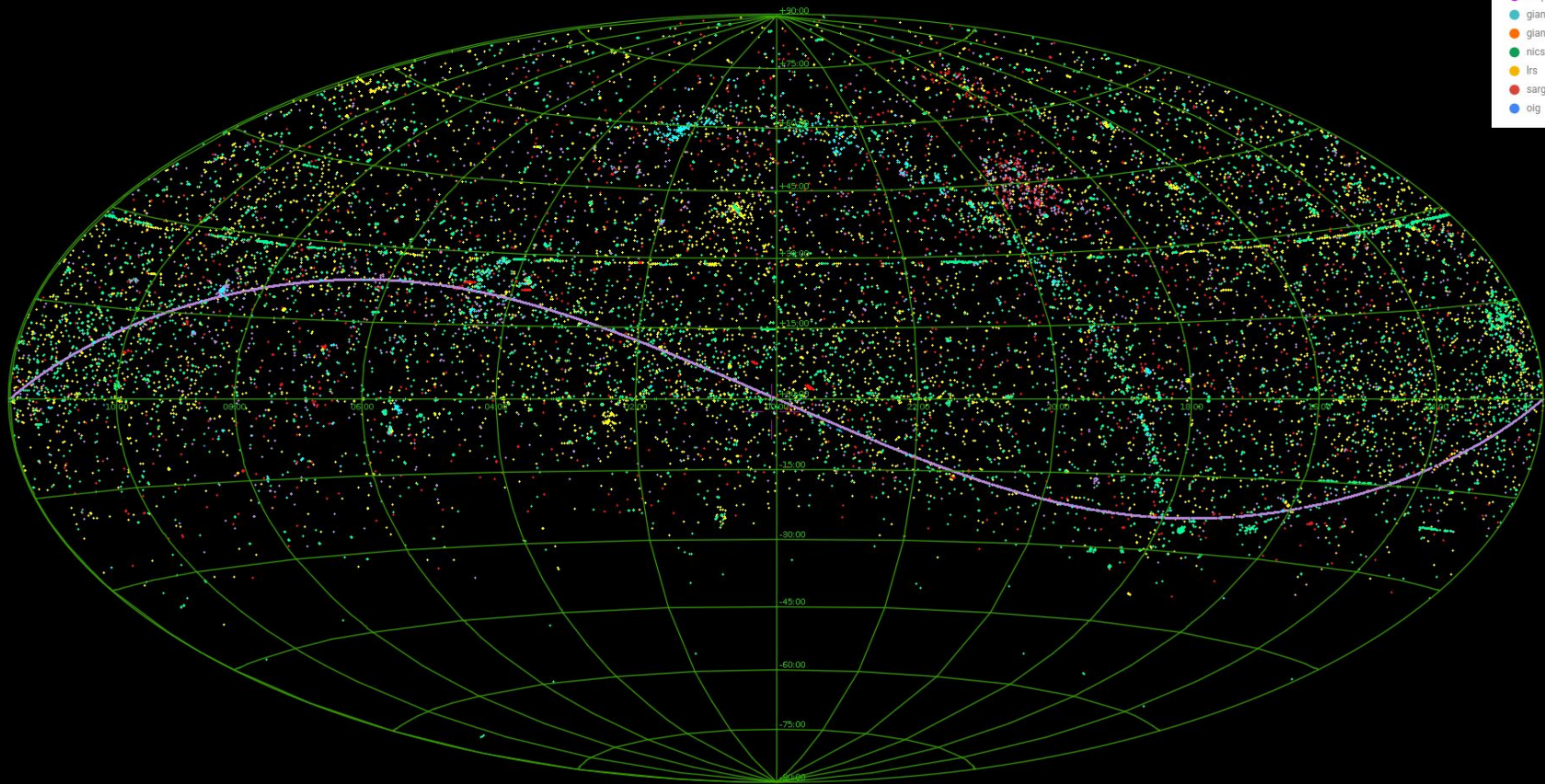
Archive Access Statistics

[portals-all] Successful actions per portal by time range long term



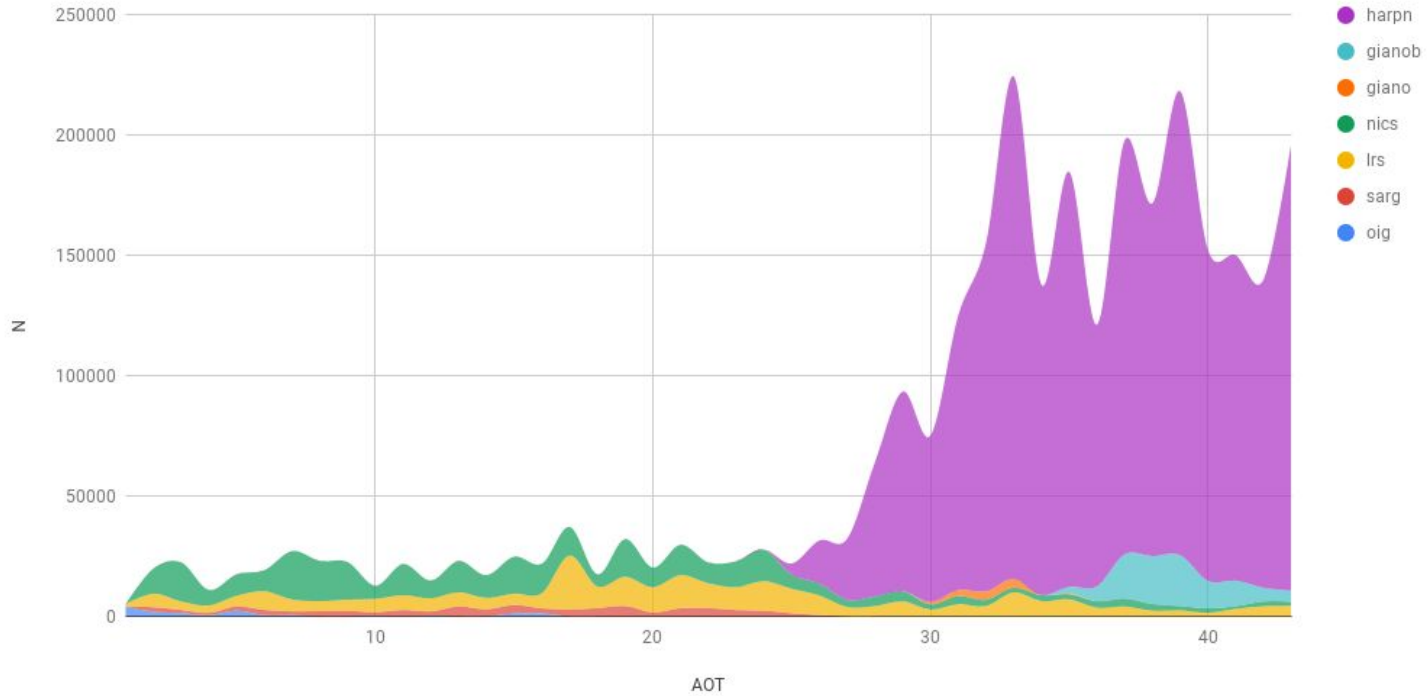
[Fileserver] Successful actions by archive long time ranges



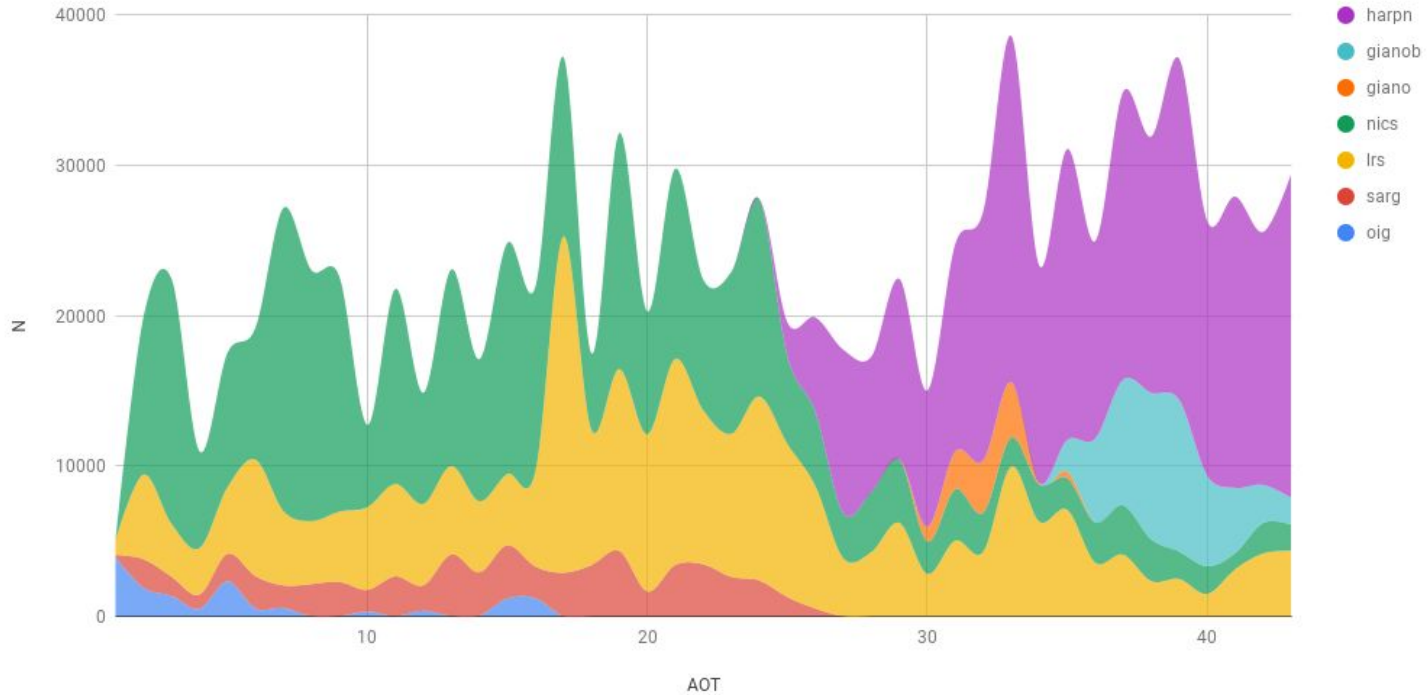


- harpn
- gianob
- gianc
- nics
- lrs
- sarg
- oig

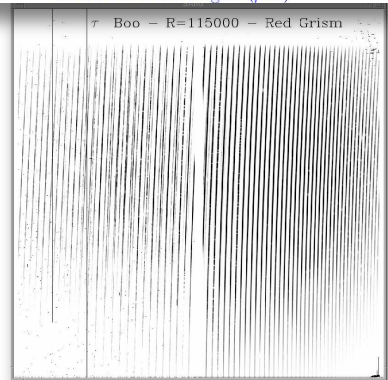
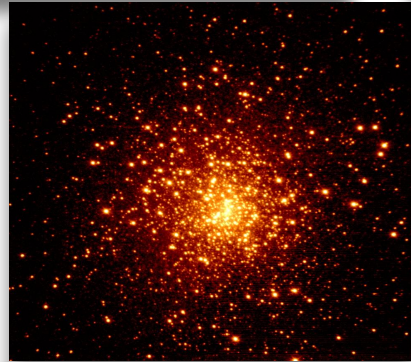
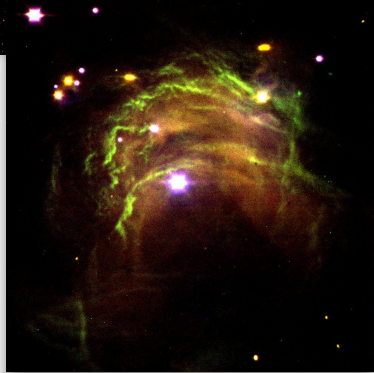
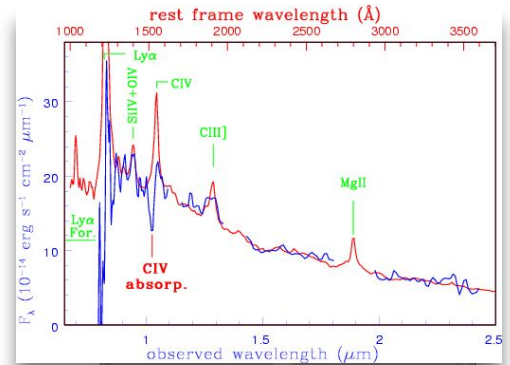
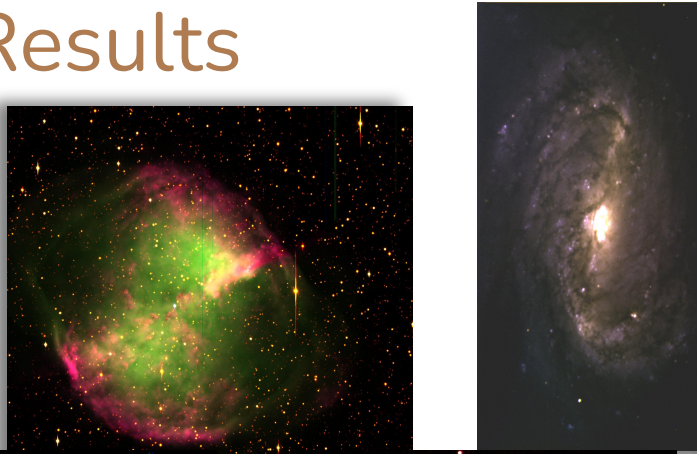
Number of Files per Instrument



Number of Raw Files per Instrument



Many Images, Spectra and Scientific Results



GAPS@IA2

GAPS (Global Architecture of Planetary Systems) is a long-term program for the comprehensive characterization of the architectural properties of planetary systems as a function of the hosts' characteristics (mass, metallicity, environment):

- more than 80 INAF and associated scientists in Italy, and from foreign institutes
- more than 23.000 spectra at TNG since August 2012

GAPS required a strong interaction with its private data

- Customizable re-process of GAPS data with appropriate spectral line mask and options;
- Perform queries on additional metadata content;
- A flexible and collaborative tool to manage additional info about the project and the observations;
- A repository where to access, synchronize, share and search for interesting data.

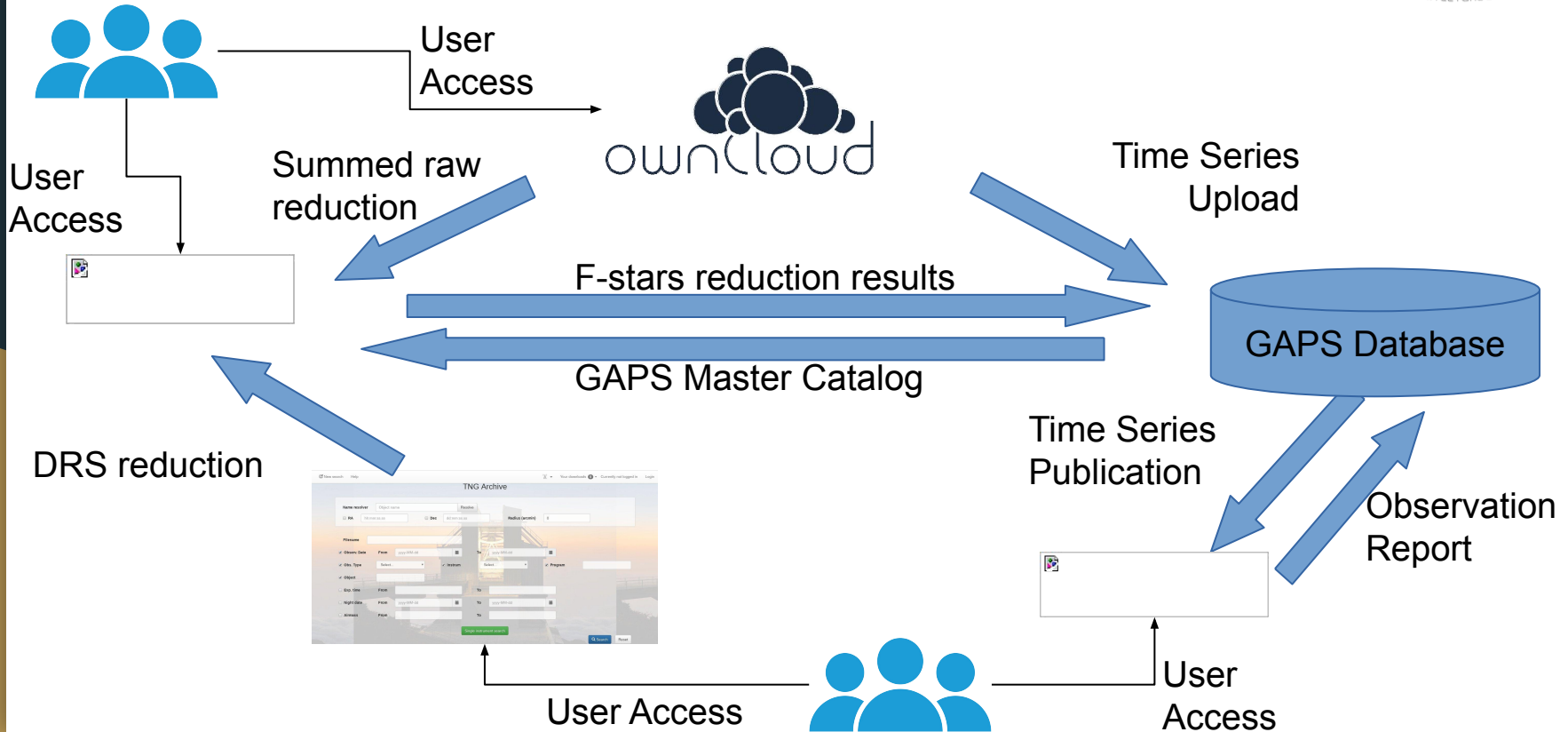
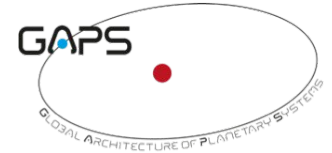
GAPS@IA2

Scientific results are not only raw or reduced data:

- Observation reports
- Time series
- Catalogs
- New masks
- Alternative reduced data
- Summed raw files
- Scientific and ancillary data shared

Solution is a custom framework of services managed by IA2 and content managed by GAPS members

GAPS@IA2



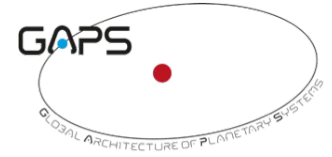
Yabi

- Yabi is a 3-tier application stack to provide users with an intuitive, easy to use, abstraction of compute and data environments. Developed at the Centre for Comparative Genomics and Murdoch University, Yabi has been deployed across a diverse set of scientific disciplines and high performance computing environments
- Yabi deployed at IA2 allows users to run custom data reduction with the HARPS-N DRS pipeline
- Raw data are automatically reduced in Trieste and saved on the Yabi storage (available 6 hours after end of observations)
- In the Yabi interface directories TNG users can find:
 - Reduced data subdivided by observation night and targets
 - Proprietary custom masks and tools



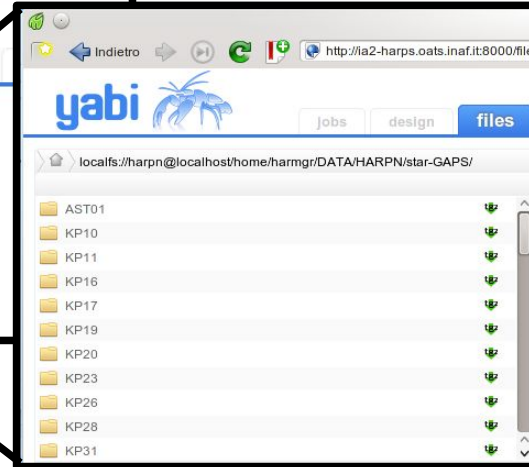
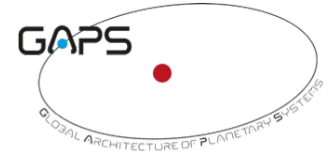
<https://ccgapps.com.au/yabi/>

Yabi

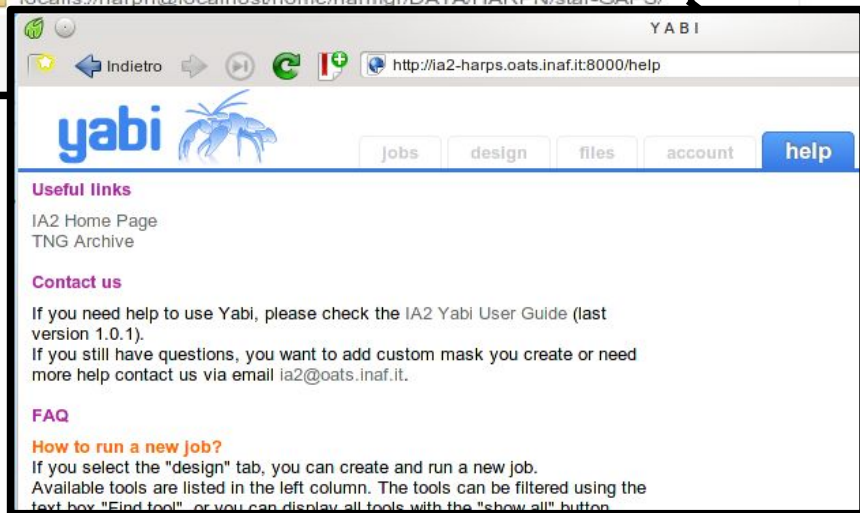
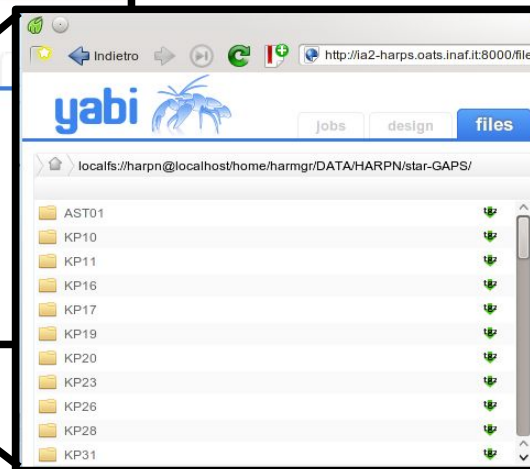
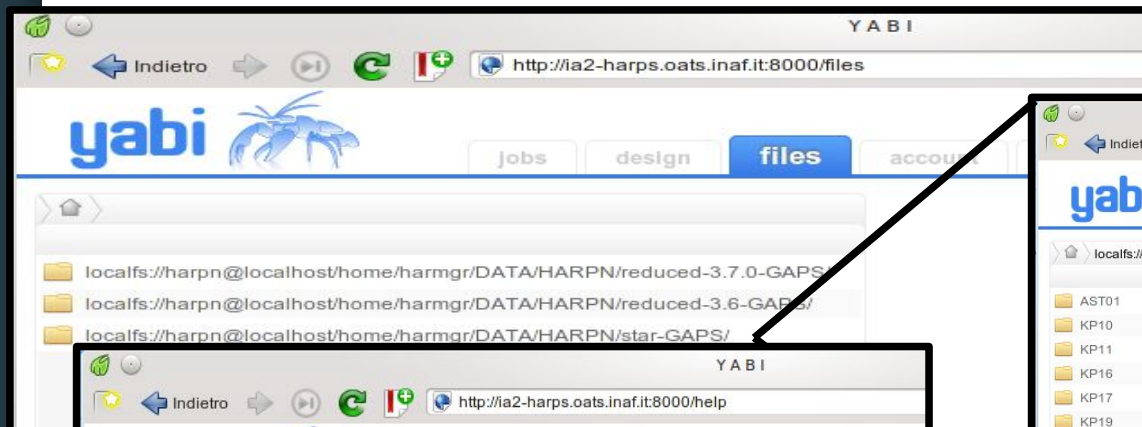
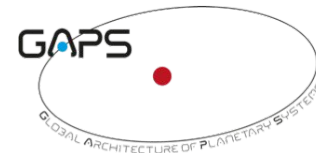
A screenshot of a web browser displaying the Yabi interface. The browser's address bar shows the URL "http://ia2-harps.oats.inaf.it:8000/files". The Yabi logo, featuring a blue crab-like creature, is visible on the left. A navigation menu includes buttons for "jobs", "design", "files" (which is highlighted), "account", and "help". The main content area shows a file browser view with a list of three directories:

- localfs://harpn@localhost/home/harmgr/DATA/HARPN/reduced-3.7.0-GAPS/
- localfs://harpn@localhost/home/harmgr/DATA/HARPN/reduced-3.6-GAPS/
- localfs://harpn@localhost/home/harmgr/DATA/HARPN/star-GAPS/

Yabi



Yabi



Yabi

- From the “design tab” add a select tool to the workflow (e.g. select target)

The screenshot shows a web browser window with the URL `http://ia2-harps.oats.inaf.it:8000/design`. The page title is "YABI". The browser's address bar contains "Indietro", "Duck Duck Go", and a copyright notice: "© 2006-2014 CCG, Murdoch University". A "log out harpn" button is visible in the top right.

The main interface has a navigation bar with tabs: "jobs", "design" (selected), "files", "account", and "help". The "design" tab is active, showing a workflow editor. On the left, there is a "file selector" panel with a "Find tool: select" dropdown and a "show all" button. Below it, a checkbox "Use selection to auto-filter?" is checked and labeled "on". The file selector contains three items: "select single e2ds file", "select sum file", and "select target", each with an "add" button.

The main workflow area is titled "Test CCF Target" and has a "run" button. It features a "Tags:" dropdown menu. The workflow is bounded by "start" and "end" labels. A large blue arrow points to a central box containing the text: "drag tools here to begin (or use the add buttons)".

Yabi

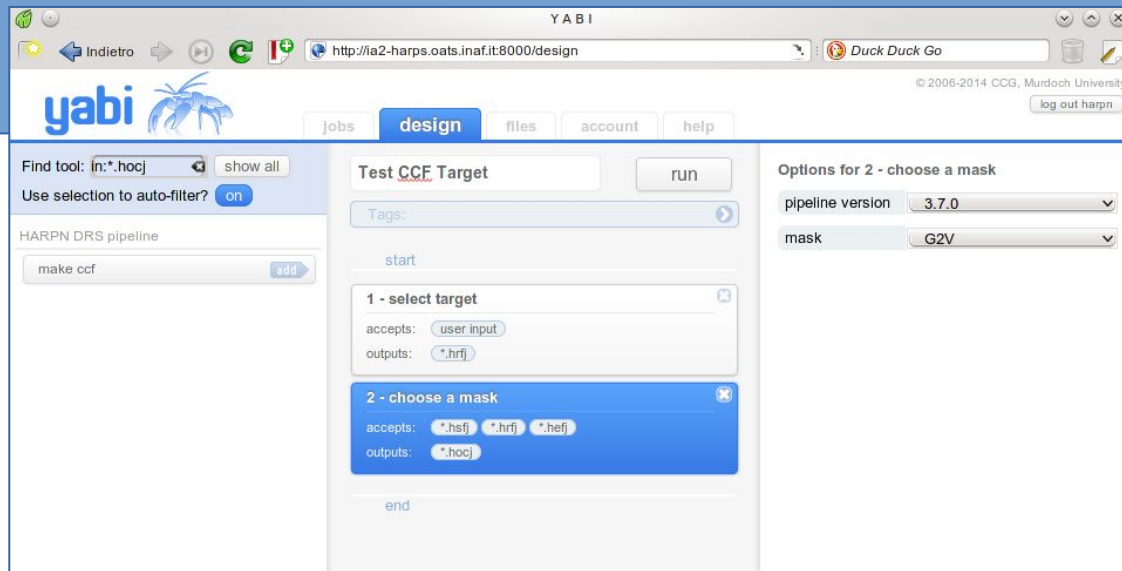
- From the “design tab” add a select tool to the workflow (e.g. select target)
- Automatically Yabi suggests the next tool to add: choose a mask

The screenshot displays the Yabi web interface in a browser window. The address bar shows the URL `http://ia2-harps.oats.inaf.it:8000/design`. The page features a navigation menu with tabs for 'jobs', 'design' (which is active), 'files', 'account', and 'help'. A 'log out harpn' button is visible in the top right corner. The main content area is divided into three sections:

- Left Panel:** Contains a search bar with the text 'Find tool: in:*.hrfj' and a 'show all' button. Below it is a checkbox for 'Use selection to auto-filter?' which is currently 'on'. A section titled 'HARPN DRS pipeline' includes a 'choose a mask' button with an 'add' button next to it.
- Center Panel:** Displays a workflow titled 'Test CCF Target' with a 'run' button. The workflow steps are: 'start', '1 - select target' (highlighted in blue), and 'end'. The '1 - select target' step shows 'accepts: user input' and 'outputs: *.hrfj'.
- Right Panel:** Titled 'Options for 1 - select target' with a 'show all options' button. It contains a 'target name' field with the value 'MP14'.

Yabi

- From the “design tab” add a select tool to the workflow (e.g. select target)
- Automatically Yabi suggests the next tool to add: choose a mask
- Finally, add the tool “make ccf” and set the last input parameters



The screenshot shows the Yabi web interface in a browser window. The address bar displays `http://ia2-harps.oats.inaf.it:8000/design`. The page has a navigation menu with tabs for 'jobs', 'design' (selected), 'files', 'account', and 'help'. A 'log out harpn' button is visible in the top right.

The main content area is titled 'Test CCF Target' and includes a 'run' button. It shows a workflow with the following steps:

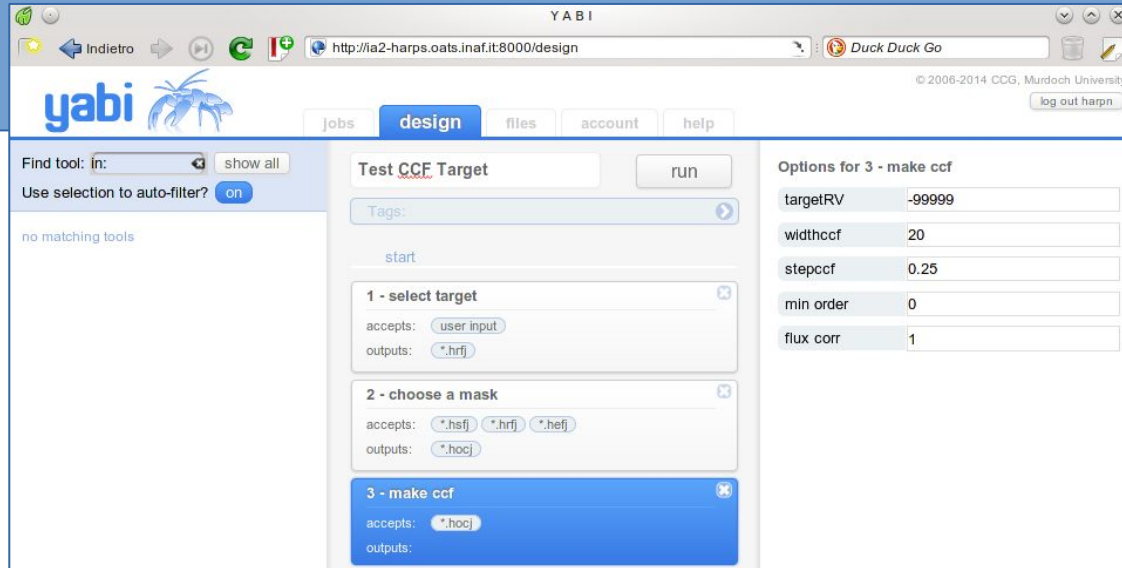
- start**
- 1 - select target**: accepts 'user input', outputs '*.hrfj'
- 2 - choose a mask** (highlighted in blue): accepts '*.hsfj', '*.hrfj', '*.hefj', outputs '*.hocj'
- end**

On the right side, 'Options for 2 - choose a mask' are shown as dropdown menus: 'pipeline version' is set to '3.7.0' and 'mask' is set to 'G2V'.

On the left side, there is a search bar for tools with the text 'Find tool: in:*.hocj' and a 'show all' button. Below it, a 'Use selection to auto-filter?' toggle is set to 'on'. A 'HARPN DRS pipeline' section contains a 'make ccf' tool with an 'add' button.

Yabi

- From the “design tab” add a select tool to the workflow (e.g. select target)
- Automatically Yabi suggests the next tool to add: choose a mask
- Finally, add the tool “make ccf” and set the last input parameters



The screenshot shows the Yabi web interface in a browser window. The address bar displays `http://ia2-harps.oats.inaf.it:8000/design`. The page has a navigation menu with tabs for 'jobs', 'design' (selected), 'files', 'account', and 'help'. A 'log out harpn' button is visible in the top right.

The main content area is titled 'Test CCF Target' and features a 'run' button. Below the title, there is a 'Tags:' field and a 'start' button. The workflow consists of three steps:

- 1 - select target**: accepts 'user input', outputs '*.hrfj'.
- 2 - choose a mask**: accepts '*.hsfj', '*.hrfj', '*.hefj', outputs '*.hocj'.
- 3 - make ccf**: accepts '*.hocj', outputs: (empty).

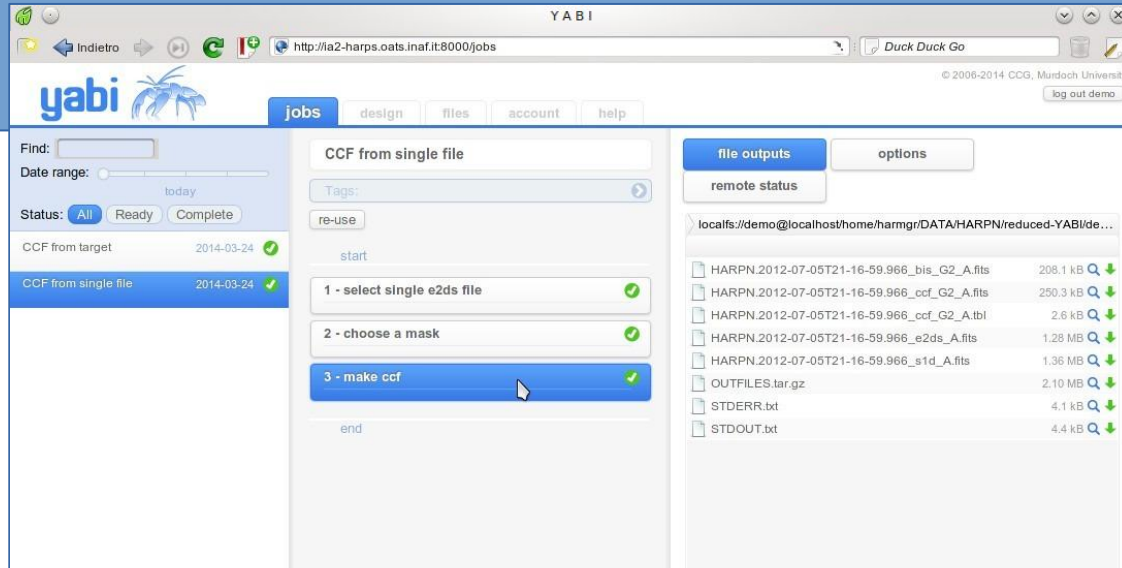
On the right side, there are 'Options for 3 - make ccf' with the following values:

targetRV	-99999
widthccf	20
stepccf	0.25
min order	0
flux corr	1

On the left side, there is a search box for tools with the text 'Find tool: in:' and a 'show all' button. Below it, it says 'Use selection to auto-filter? on'. A message indicates 'no matching tools'.

Yabi

- From the “design tab” add a select tool to the workflow (e.g. select target)
- Automatically Yabi suggests the next tool to add: choose a mask
- Finally, add the tool “make ccf” and set the last input parameters



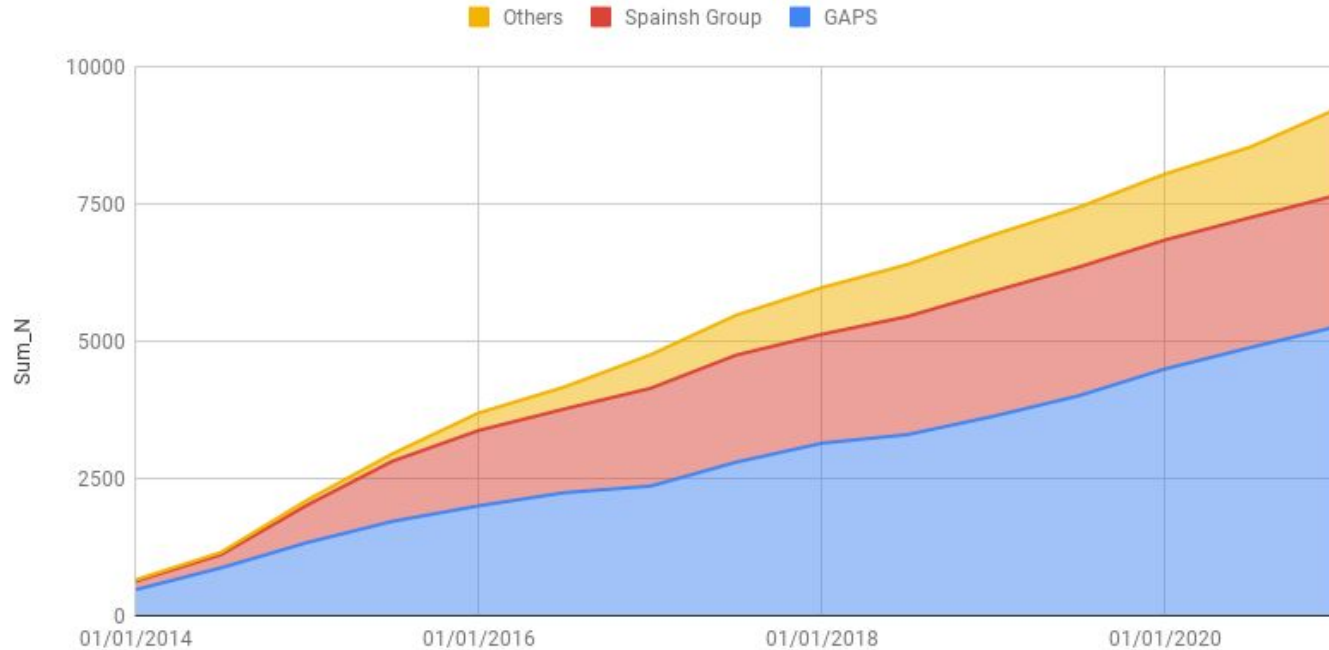
The screenshot shows the Yabi web interface in a browser window. The URL is <http://ia2-harps.oats.inaf.it:8000/jobs>. The interface has a navigation bar with tabs for 'jobs', 'design', 'files', 'account', and 'help'. The 'jobs' tab is active, showing a workflow titled 'CCF from single file'. The workflow consists of three steps: 1 - select single e2ds file, 2 - choose a mask, and 3 - make ccf. The 'make ccf' step is currently selected. On the right, there is a 'file outputs' section showing a list of files with their sizes and download links. The files listed are:

File Name	Size	Download
HARPN.2012-07-05T21-16-59.966_bis_G2_A.fits	208.1 kB	Download
HARPN.2012-07-05T21-16-59.966_ccf_G2_A.fits	250.3 kB	Download
HARPN.2012-07-05T21-16-59.966_ccf_G2_A.tbi	2.6 kB	Download
HARPN.2012-07-05T21-16-59.966_e2ds_A.fits	1.28 MB	Download
HARPN.2012-07-05T21-16-59.966_s1_d_A.fits	1.36 MB	Download
OUTFILES.tar.gz	2.10 MB	Download
STDERR.txt	4.1 kB	Download
STDOUT.txt	4.4 kB	Download

Yabi Statistics

- Yabi access extended to all HARPS-N users
 - 260 total users
 - 97 active users
- Tools and masks developed
 - 14 tools, 8 different kind of workflows
 - 38 custom masks developed by users
- 9490 workflows executed since March 2014

Yabi Workflows per User



Yabi PoC on Google Cloud and CWL

- Proof of Concept “Yabi Workflow Execution on Google Cloud Platform” to run data reduction pipelines on TNG archive data
 - Goals: simplify infrastructure management (SaaS/PaaS) and software deployment (Docker), optimizing and balancing the scalability of the service (Slurm and Kubernetes)
 - Results: excellent scalability and good costs (total estimated charges of 200 EUR/month to maintain architecture up and running on GCP)
 - Criticalities: needs to write data reduction pipeline optimized for containerization
- CWL (Common Workflow Language) integration in Yabi
 - CWL is a standard, it can provide an high level of interoperability between WMS and portability across different hardware environments
 - CWL supports natively Docker
 - CWL is not user friendly, it has a steep learning curve and it lacks a graphical interface to design workflows

Discover and Access GAPS Time Series within the VO Framework

- Increasing interest in the IVOA (International Virtual Observatory Alliance) for Time Domain data
- IA2 is developing a TAP VO service to publish GAPS public HARPS-N time series (with an eye on GIANO-B/GIARPS)
- Identified 8 use cases for RV Time Series discovery:
 - Number of points
 - Exoplanet properties
 - Discovery method
 - Host star properties
 - Photometry
- Exoplanets time series discovery and access using VO ObsCore Standard looks feasible... but
- Some information, useful when dealing with spectroscopic RV time series may be misleading
- Reference tables and dictionary needed



Conclusions

- Target of the archive:
 - Make services and maintain them
- Target of the archive user:
 - Use archive services and check their functionalities
 - Reporting problems
 - Suggest upgrade and new ideas to analyse data
- TNG archival data very fruitful for technological research at IA2 Data Centre:
 - New experience acquired to create and manage new archive services
 - Extend access to new services to the astronomical community