Italian Astronomical Archives IA2 - 2022

Cristina Knapic
24/05/2022

IA2 *Italian Center for Astronomical Archives*
Centro Italiano Archivi Astronomici

Home | Services | Projects | Software | Additional Info | IA2 Group

ABOUT US

collapse [-]

IA2 (Italian center for Astronomical Archive) is an Italian Astrophysical research e-infrastructure project that aims at co-ordinating different national initiatives to improve the quality of astrophysical data services. It aims at co-ordinating these developments and facilitating access to this data for research purposes. The IA2 is supported by INAF since 2005. IA2's main goals consist in data archiving systems and safety, including data hosting and data curation and preservation, data and metadata distribution over geographical sites, access services including publication within the VO scenario. IA2 provides also services and tool to the community, like data sharing (owncloud), project management (redmine), software collaboration (git-lab) and has available a workflow manager (Yabi) for computational needs.

contact us

MAIN ACTIVITIES

collapse [-]

TELESCOPE ARCHIVES & SIMULATIONS

MAIN: TNG, LBT, Asiago, SVAS, RADIO

MAIN: Exoclimates, INTRIGOSS, LBT OLD

HOSTED: Byurakan, INES, BaSTI, IBIS-A, other archives @ INAF

OTHER SERVICES

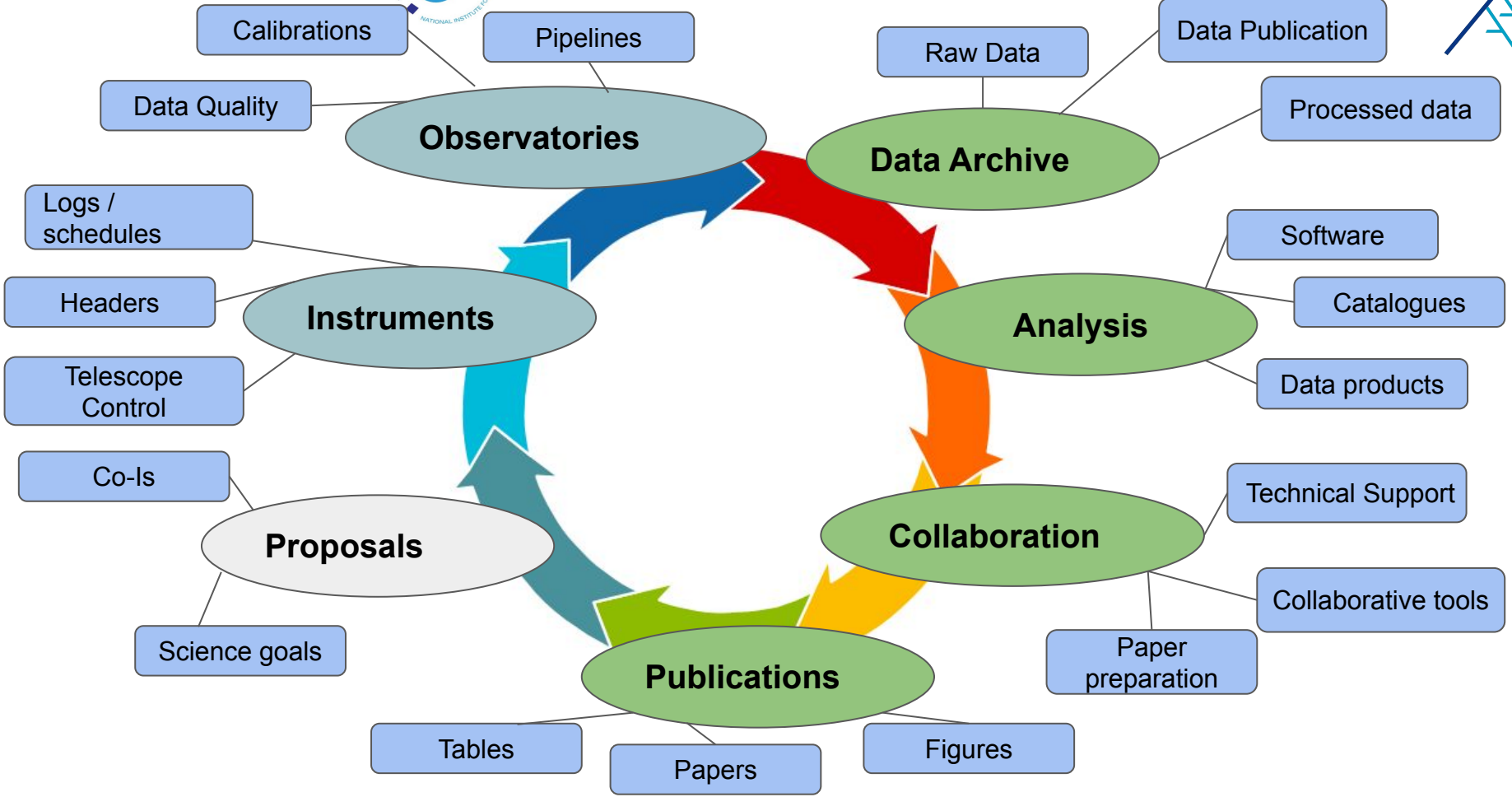
ownCloud, redmine, VO initiatives, GitLab, yabi

PROJECTS

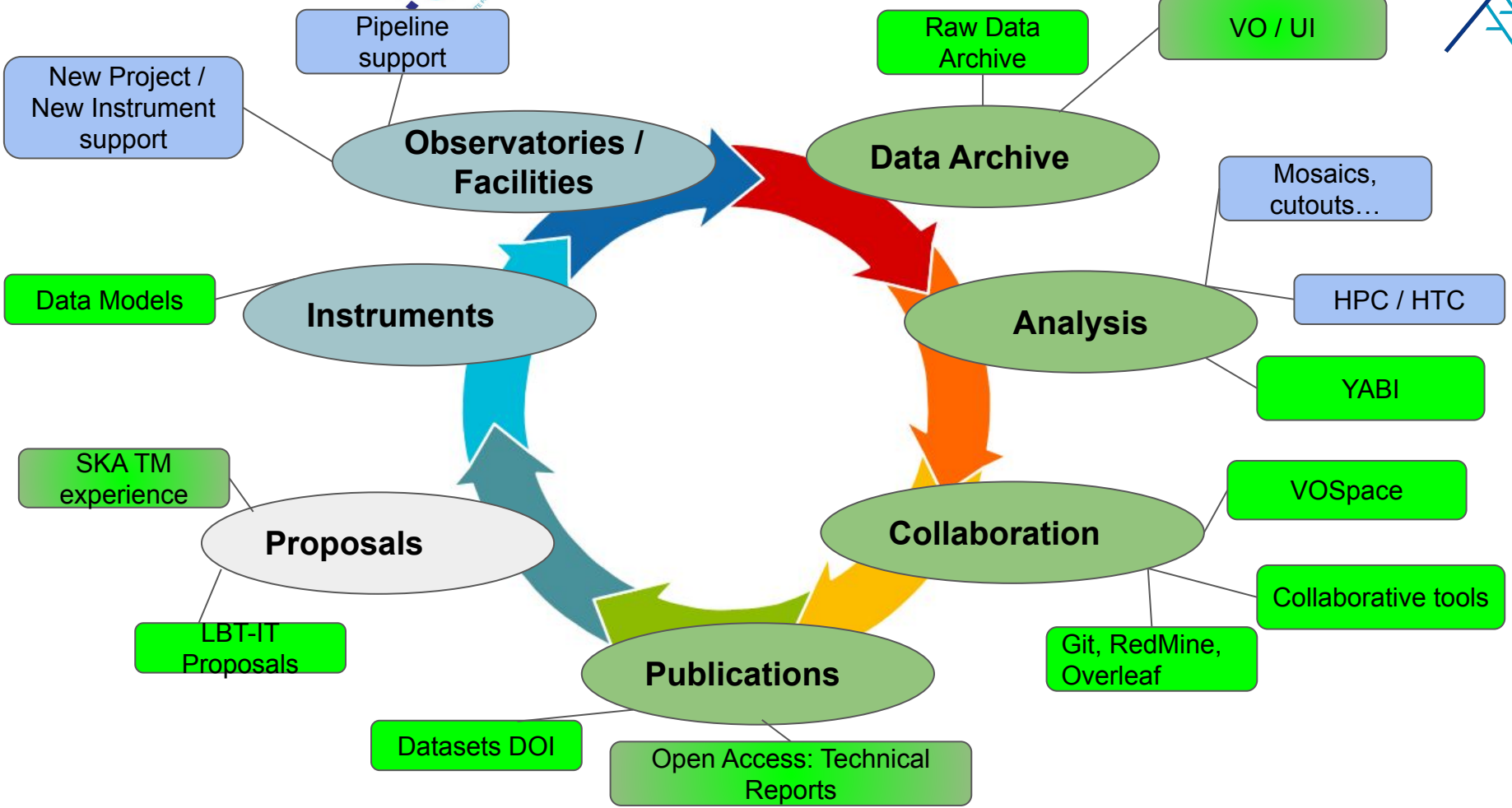
expand [+]

https://ia2.inaf.it/

Astronomical Data Life Cycle



IA2 activities:



Supported Projects

IA2 hosts data and services for several Projects:

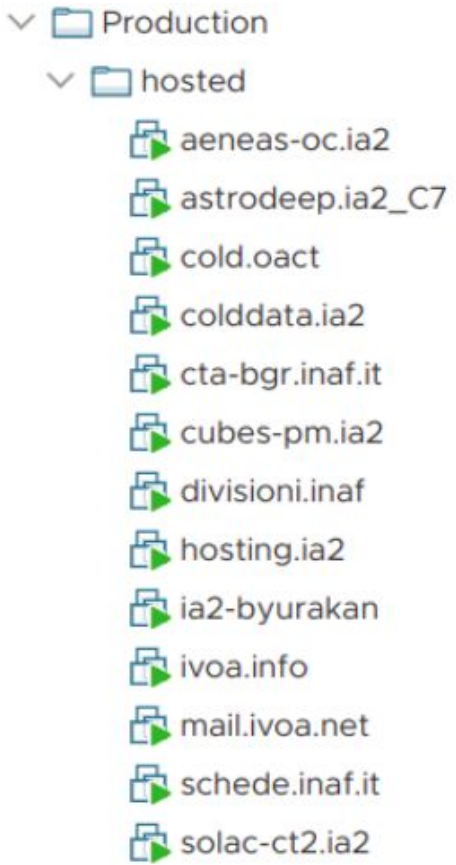
- ECOGAL;
- EUROPLANET;
- ARI-L;
- CUBES;
- MINCE;
- ViaLactea / VLKB;
- SKA DC;

It also hosts web pages and services:

- Home page of the Astronomical Observatory of Rome;
- Solar Telescope of the Catania Astrophysical Observatory;
- Schede INAF;
- Computing Catania
- Open Access storage;

IA2 activities to support INAF

- INAF survey;
- INAF researchers to store simulated data out of CINECA



Contribution to Projects

IA2 personnels were/ are involved in:

- AENEAS;
- INDIGO Data Cloud;
- SKA TM;
- SKA RC;
- PNRR;
- IVOA;
- EuroPlanet;
- ARI-L;
- ….

IA2 host also all ICT in house services

IA2 participated to

- Google platform PoC
- IBM Spectrum Scale PoC

IA2 expertise:

- data archive handling;
- database management;
- data center architecture;
- System Administration;
- web based application development;
- data distribution software development;
- Distributed and parallel filesystems (GPFS, LUSTRE);
- authentication and authorization systems;
- astronomical interoperability standards;
- services provision;
- user support;
- workflow management system & pipelines;
- scientific use cases (radio and optical)

Infrastructure: HW

Storage

LUSTRE:
600TB + 1,2PB on JBOD

Synology: 270TB
QSAN: 400TB

Transfer Node:
IBM 100TB all flash

Cross Backup
Synology +LUSTRE --> QSAN;
QSAN --> Synology

Total: 2.6 TB on-line

Virtualization
3 SuperMicro server 40CPU 512GB RAM
2 Lenovo 32 CPU 1 TB RAM

Storage for VMs
100TB all flash
10TB on HD

DB machines
2 servers DELL 2CPU (2x32 core)
256GB RAM
3.5 TB all flash

Total: 90 on-line VMs

Infrastructure: HW

TS450 - 6 drive LTO7

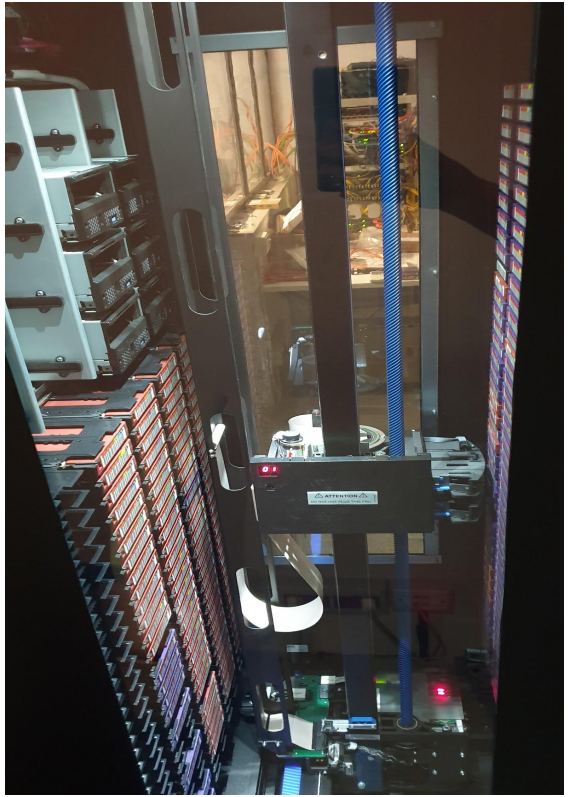
240 Cartridge of ~ 9GB

Total: **~2PB**

Data in Tape:
400 TB used (78 cartridge)
1600TB free (162 cartridge)
Free space is not real since we must check data before copying them on Tape (see ARI-L project).

Front-end: Lenovo server 92 TB all flash
70 TB licensed IBM Spectrum Scale and Archive (File System)
Partitions:

```
Filesystem              Size  Used Avail Use% Mounted on
lv_archive_01            25T   24G   25T   1% /ia2_tape_archive_01
lv_generic_rw_01         13T  3.1T  9.2T  26% /ia2_tape_generic_rw_01
lv_stb_01                32T   14T   18T  44% /ia2_tape_stb_01
```



Infrastructure: HW

- Future expansions:
  - Tape library:
    - Completion of the LTO7 rack;
    - LTO8 new rack;
  - Transfer node expansion to full capacity all flash (+ 100TB)
  - Virtualization system:
    - acquisition of at least 2 servers (2x32 core; 1TB RAM)
    - storage expansion all flash (+100TB)
- Net: request for a data dedicated subnet through INAF data centers.

Infrastructure: SW

- **ICT** services (on IA2 infrastructure)
  - 280/day di accesso alla home page ICT
  - **owncloud** (130 utenti)
  - **redmine** (~250 Utenti)
  - **gitlab** (~250 Utenti / ~ 370 progetti)
  - **indico** (~ 1100 inaf; ~700 esterni)
  - **DOI** service https://doi.ict.inaf.it/

Infrastructure: SW

Virtualization: VMWARE

Monitoring: Ganglia & Nagios

Logging: Kibana on db

License server

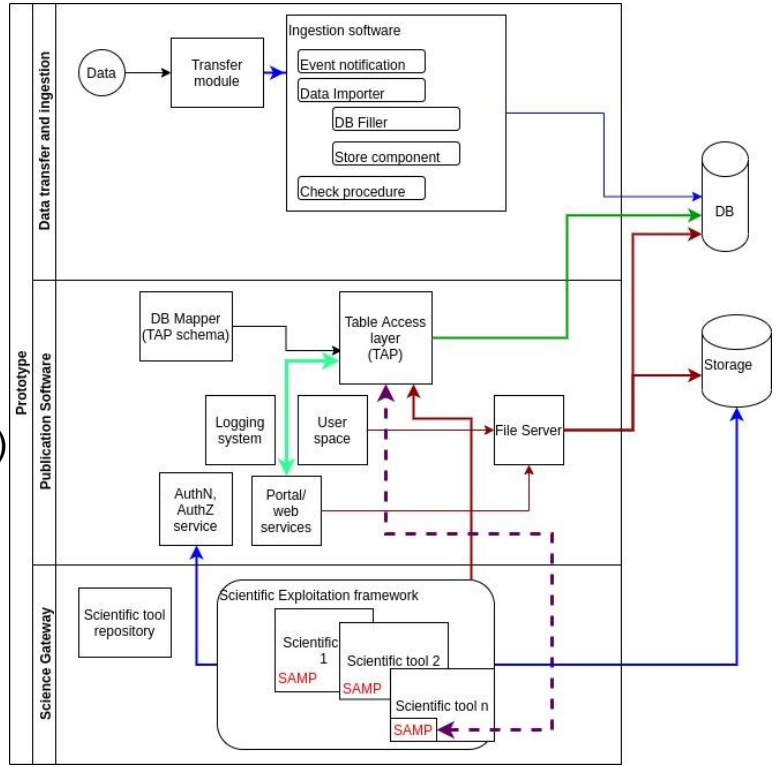
Infrastructure: SW

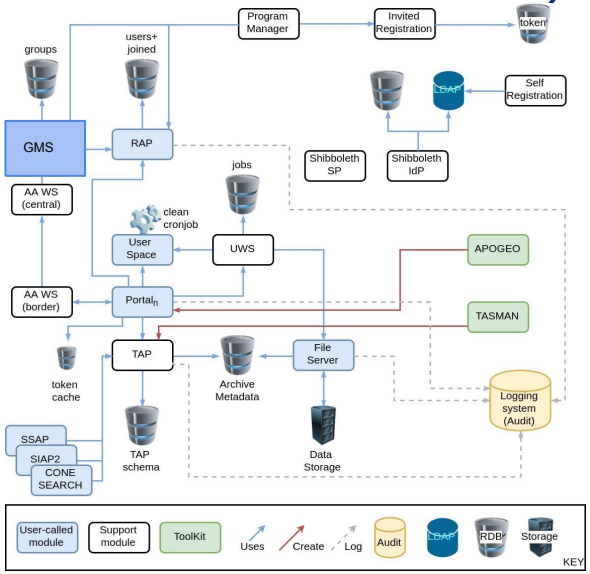
Developed Software:

NADIR
NEXT (RDI)
Nadir2PG
Web Portals
Archive Program Manager
Exposure Time Calculator
Proposal web tool
TAP lib
GMS service
SSO system (RAP + GMS)
File Service
APOGEO
TASMAN
VOSSIA

and

VOSpace





Infrastructure: SW

The INAF VOSpace implementation

https://drive.google.com/file/d/1BqYOqImdH5J6LOqO81suomOf-dZv1vou/view

Infrastructure: SW

VOSpace next feature will be:

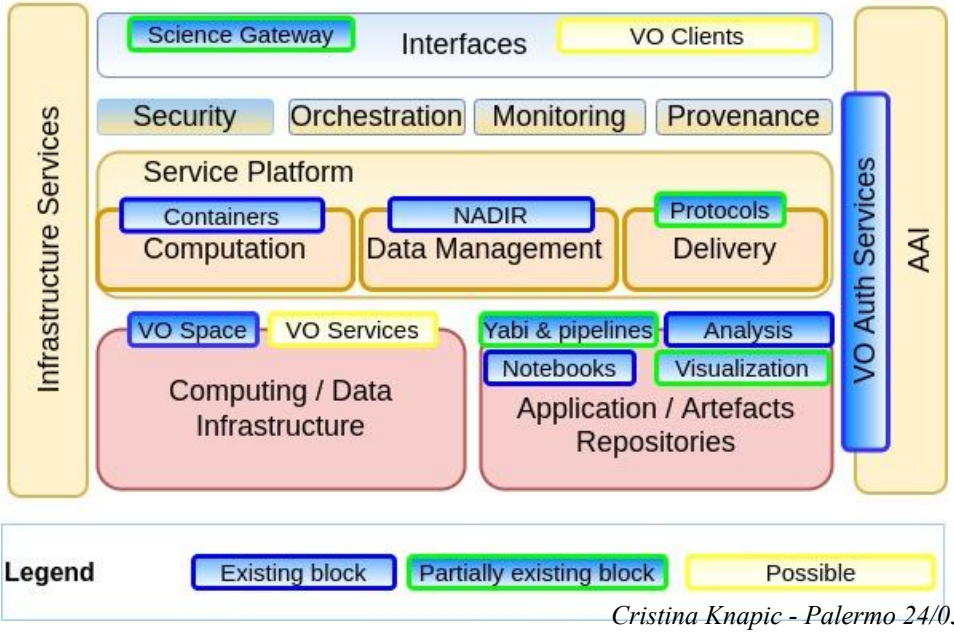
- **DOI integration.** It will allow users to request for a dataset DOI without duplication of data (backup is anyway applied);
- direct **connection to computational clusters** (Pleiadi);
- **ssh** connection to VOSpace (moderated by admin) for big data storage (Carmelita use case);

IA2 next features:

- **backup** (geographical distribution) and **integrity** checking strategies;
- git (in support to ICT)
  - **CD/CI** (already internally experimented)
  - availability of a **runner** machine (planned in the scope of ALMAGAL preliminary tests)
- ssh storage access using the IDEM credentials (no possibility to perform other actions)
- **vms/containerization** (in support to ICT)
  - availability of a limited number of VMs for projects that require a web portal or limited computational resources;
- **YABI** : more space for accessing pipelines
- Archive meta-description enrichment with **references to published papers**.

IA2 is ready to be part of a INAF Science Gateway

Infrastructure: SW



Several blocks are already in place, some others are under development.

All we need is … a science gateway to collect them all!

Thanks!

Questions? Hints?