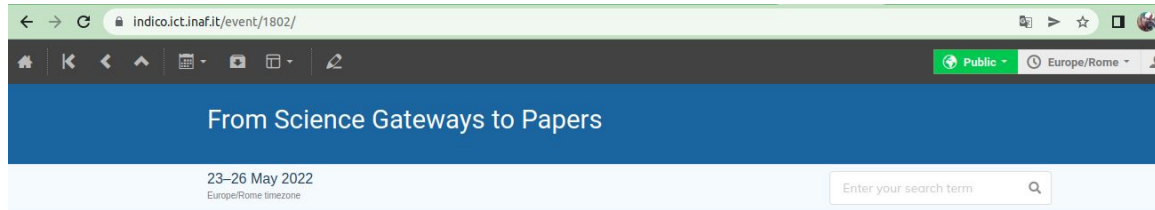


# From Science Gateway to Papers - 2022



Introduzione
Agenda
Contribution List
My Conference
<b>My Contributions</b>
Registrazione
Call for Abstracts
Lista dei partecipanti
Informazioni logistiche
Descrizione delle sessioni
<b>Dr. Francesca Martines</b>
✉ <a href="mailto:francesca.martines@inaf.it">francesca.martines@inaf.it</a>
✉ <a href="mailto:cristina.knopic@inaf.it">cristina.knopic@inaf.it</a>



I temi caldi dello sviluppo tecnologico di supporto alla comunità scientifica sono oggi centrati sulla offerta di Science Gateway di facile utilizzo, ove siano presenti i dati ed anche tools di analisi e visualizzazione. Negli ultimi anni sono state avviate esperienze di questo tipo in ambiti anche molto diversi tra loro, e appare quindi il momento di confrontare le esperienze acquisite per identificare future roadmap di sviluppo. Questo Workshop si pone quindi l'obiettivo di fare il punto sulle reali necessità

*Cristina Knopic  
on behalf of SOC  
24/05/2022*

# Goals



“The purpose of the workshop is to gather for discussion all main Italian actors involved in the use and management of astrophysical data - observed or simulated-, also within the perspective of multimessenger and multidisciplinary sources. From an overview of the existing science gateways and their development, to the discussion of the science gateway concept for the Big Data, different existing framework and tools will be presented and discussed. The use of modern era data is no longer circumscribed to the search for scientific information, but it extends to provide the framework for the search, manipulation, analysis and publications of data and papers from telescopes, either terrestrial or satellite, of the new telescope era. Therefore, hints and requirements from the community is highly stimulated and appreciated in order to improve the institutional offered services to support Science”.



# Participants and organization

## SOC:

- Cristina Knapic (chair)
- Ugo Becciani
- Marcello Lodi
- Francesca Martines
- Marco Scodeggio
- A. Zanichelli

## LOC:

- Francesca Martines
- Laura Daricello
- Laura Leonardi
- Giulia Manca
- Flavio Morale
- Salvatore Speciale

## Invited Speakers:

- Francesco Bedosti
- Jan Brand
- Vincenzo Galluzzi
- Adriana Gargiulo
- Uta GrohtKopt
- Avet Harutyunyan
- Marco Landoni
- Paola Masuzzo
- Eva Sciacca
- Franco Tinarelli

*Si ringrazia l'Osservatorio di Palermo per la logistica e il supporto*

# Session topics



## Session 1 - Example of Science Gateways in International Infrastructures

We will illustrate and discuss the characteristics of the major Science Gateways existing at the international level that support large-scale scientific projects, not necessarily related exclusively to astrophysics. What exists and what is useful for our science.

## Session 2 - INAF Infrastructures to support Science

We want to illustrate and discuss the implementations that we have developed in our institution for both large national and international and general purpose projects and we want to discuss the effectiveness of the solutions offered by comparing them with the needs of the reference scientific community.

Followed by a round table on the use and available features.

## Session 3/4 - Examples of Science Gateways applied to projects

The major contributions to the definition of SG will be discussed in the context of major European projects with particular attention to the results and suggested architectures.

Followed by a round table on possible implementations.

# Session topics



## Session 5 - HW and SW infrastructure

Session dedicated to current INAF implementations with particular regard to hardware infrastructures and software developments and installations for their management (containerization / virtualization / packages). Use cases and potential hints for the Institute will be presented and discussed.

## Session 6 - New / starting infrastructures

We want to give an overview of what will soon be available by focusing on customization in view of specific use for projects or observational infrastructures (eg. PON SRT or TECNOPOLO). Following free space for questions and insights.

## Sessione 7 - Pipelines and their management

We want to analyze past experiences by comparing them with the new needs for data reduction, illustrate the potential of remote computing and the use of remote servers for the reduction and analysis of observational / astrophysical simulation data, discussing the modus operandi, the logics of access and the advantages offered by this approach.

## Sessione 8 - SG tools and applications

We want to discuss the most recent strategies for the management of computing systems (containerization / virtualization / packages) and the major applications for sharing data products and documentation relating to their production through, for example, the use of Notebooks annotation tools or collaborative tools at large.

# Session topics



## Session 9 - FAIRness, meaning and virtuous examples

Following the suggestions of the European community for the management of Open Science, we want to give definitions and suggestions on how to operate in order to guarantee the FAIRness of scientific data and publications. Clarification questions from the audience are highly desirable.

## Session 10 - FAIR Systems

Details of the use and implementation of the major international data storage systems (ESO, IDIA) connected to the bibliographic part of the membership will be introduced.

## Session 11 - Proposals for INAF

Solutions and proposals to meet the FAIRness needs of the institution, with a round table on the ease of finding the information necessary to achieve FAIRness and use of the proposed tools.

## Session 12 - INAF policy on Research products

Discussion on policy, intellectual property, copyright and management of embargo times of observational data and research products in INAF. Tips for achieving FAIRness.

# Round tables



During the numerous round tables slots, large space is left to suggestions coming from the scientific community since several new requirements and needs emerged in the last years, i.e.

- Facilities, projects and collaborations;
- archived data close to computational facilities: the net;
- containers, VMs and packages: how to use them;
- interactive (Big!) data analysis and visualization;
- collaborative environments and tools to allow products finalization;
- science ready data becoming Open and FAIR;
- DOI request from publishers;
- increase data citations

All these can be done and needs your collaboration!

# Aside issues



- Workflow management systems and pipelines
  - Pipelines are instrument specific and often not publicly available.
    - Pipeline accessibility ?
    - Data quality management ?
    - Super users to handle pipelines ?
  - functionalities
    - privacy of pipelines;
    - interactive workflow;
    - software stack;
- products handling
  - policy and intellectual property;
  - accessibility of the data products (Open Science and DOI)
- Data Management Plans: the lost chapter





# Restart from where we left



- Future solutions:
  - Visione della Direzione Scientifica:
    - “INAF astronomical community coming late with respect to other communities (f.i. genomics) wrt big data, HPC, archiving. What we need is mentality, effort and infrastructure” -> discussione con gli invited sulle facilities internazionali
    - “New figures in INAF are needed, who are in the middle between science and technology..” -> discussione sulle figure a supporto e di supporto dei centri dati/computing;
  - Cosa si fa in Europa: EOSC -> not for free! E i dati devono essere curati!!
  - Preservazione: DOI e Open Science, il futuro delle nostre pubblicazioni. Abbiamo il servizio DOI ma come possiamo massimizzare l'utilizzo e il problema della proprietà intellettuale delle pubblicazioni su rivista o meno;
  - Importanza dell'interoperabilità -> esistono degli standard e delle best practices, usiamole!
    - il concetto di data curation: si estende dalla conservazione dell'accessibilità al dato e al fatto che i formati divengono desueti al completamento delle informazioni corollarie (dalla data quality alla provenance..)
    - meta descrittori e WCS sono concetti trascurati perchè non si ha la visione del riuso;
  - Modelli, formati e servizi: una catena frammentata che deve diventare un tutt'uno!

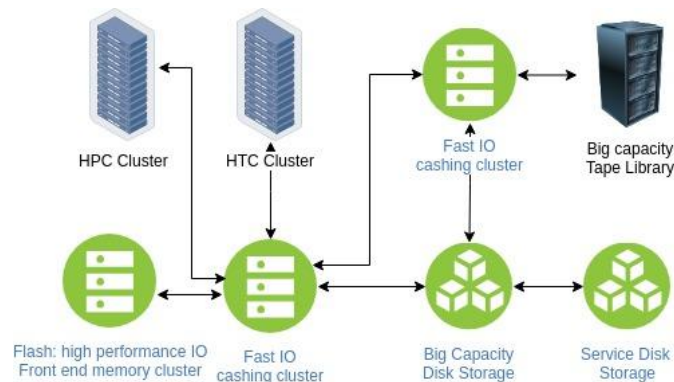
# Restart from where we left

- Science Gateways:
  - Non solo experiment devoted ma sistemi in grado di permettere un reale riutilizzo del contenuto scientifico dei dati acquisiti e fluidi a sufficienza da garantire un incremento graduale delle risorse a disposizione. Lo science gateway è visto come un catalizzatore di risorse (dati, strumenti e applicazioni) a supporto della comprensione scientifica del caso.
    - strumenti: computazione, storage, supporto...
    - applicazioni: analisi, visualizzazione, calibrazione, cooperazione, servizi di archivio, interoperabilità...
    - dati: ricerca, accesso, produzione, condivisione...
  - Aggancio della parte archivi con le necessità di calcolo;
  - Il calcolo ai dati: in prospettiva di grandi infrastrutture osservative internazionali, è ragionevole pensare di ottimizzare i tempi scala di processamento avanzato presso i data center spostando la computazione verso i dati. Questo implica un cambiamento di prospettiva per chi sviluppa gli science gateway e soprattutto di chi li utilizzerà.
    - USER SPACE:
      - non più archivio da cui scaricare in locale dati, ma ambiente integrato con la possibilità di fare calcolo, applicare algoritmi, trovare tools per l'analisi;
    - SUPPORTO: la complessità di alcuni progetti implica la necessità di avere supporto sia scientifico che informatico per la corretta riduzione ed analisi dei dati.

# Restart from where we left



- Archivi e computazione (HW):
  - Archivi:
    - multi-tiered;
    - distribuiti;
    - business logic coordinata e interoperabile;
  - Computazione:
    - nuove tecnologie (esempio di LOFAR): quanto sono facilmente utilizzabili?
    - la mentalità vs cambio approccio;
    - i PoC con i provider cloud esterni come sono andati?
    - è comunque indispensabile il supporto scientifico tecnico;



# Enjoy the meeting!!

