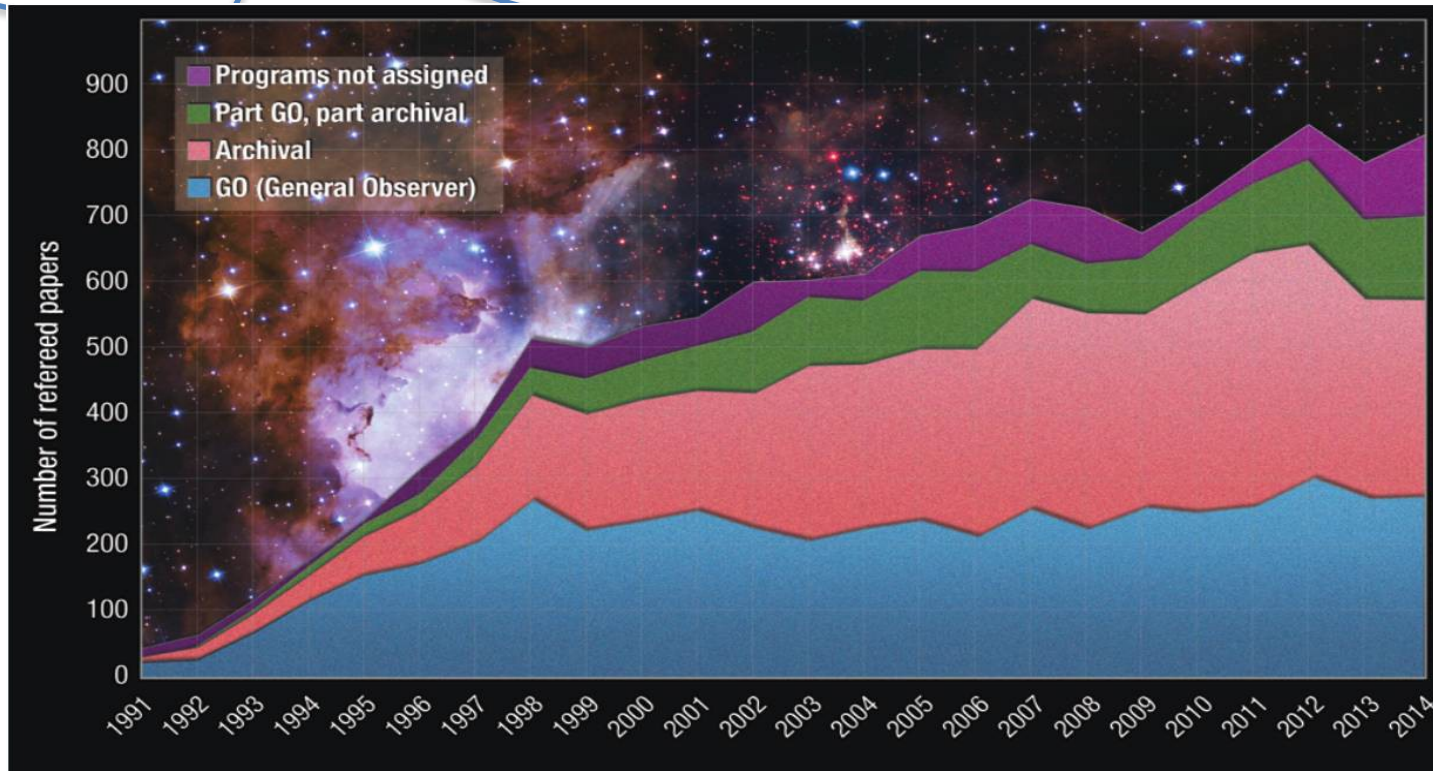


# INAF Strategy in the Big Data Era

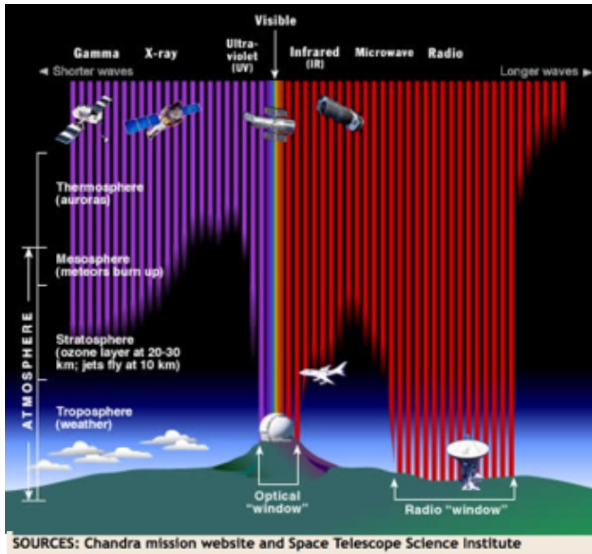
R. Smareglia

Head of ICT and Data Management office of INAF

GARR Workshop, Venezia 15-17 Novembre 2017



HST Newsletter: “At the present time, approximately **half of the refereed publications** based on Hubble observations are derived purely **from archival data**, and, every year, this number is slightly higher than the number of publications based on new observations. .... the Hubble Archive has become a goldmine for the astronomical community...”

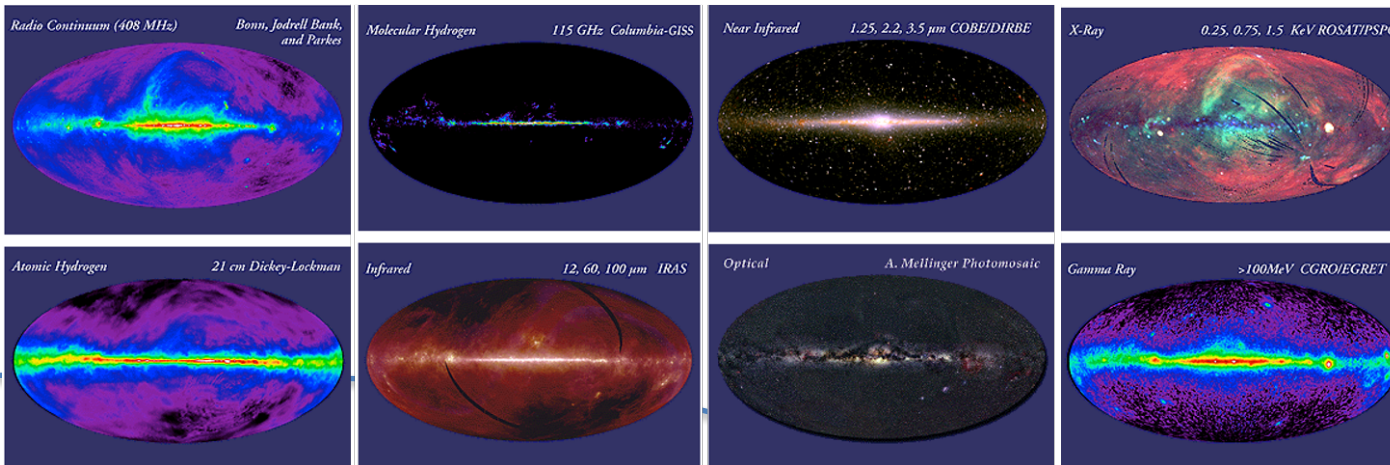


Una visione parziale delle osservazioni astronomiche nei prossimi 10 anni

Alla lista poi mancano:

- Diversi osservatori
- Le missioni planetarie
- Esopianeti
- Osservazione del Sole
- Cosmic ray exp.

	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
<b>Low Frequency Radio</b>											
LOFAR											
MWA											
VLITE on ASKA											
<b>Low/Mid Frequency Radio</b>											
SKA 1st											
<b>Mid-High Frequency Radio</b>											
AASKAP											
MAT → MURAT											
PIA											
SMRT											
RTCA											
(sub) Millimeter Radio											
ALMA											
IRRT											
<b>Optical Transient Factories/Transient Finders</b>											
PTF											
Pan-STARRS PS1 → Pan-STARRS PS2											
Blueshift Bleritch → full array in Oct 2019											
LSST											
<b>Optical/IR Large Facilities</b>											
ESA/GAIA											
ESOS/NT											
Keck											
ESO/VST (survey)											
ESO/NST (survey)											
NASA/WFIRST											
DMT											
ESA/Euclid											
ESO/ELT											
TMT											
<b>X-ray</b>											
NASA/Swift											
ESA/XMM											
NASA/Chandra											
NASA/NuSTAR											
IBD/ASTROSAT											
Astrol											
NASA/DMT											
ATHENA (2028)											
<b>Gamma-Ray</b>											
ESA/Integral											
ASI/AGILE											
NASA/Fermi											
SWGO											
DMPE											
CTA Construction											
CTA Early science											
CTA Full Operation											
<b>Gravitational Waves</b>											
Advanced VIRGO											
Advanced LIGO											
KAGRA											
<b>Neutrinos</b>											
IceCube											
ASTARIS											
KAMNET-1											
KAMNET-2											



# Richiesta Principale: Interoperability → VO → EOSC

## Distributed resources

- International team members can bring regional resources
- Big data: moving code to data
- Resources are not simple

## Science teams

- Science teams are international virtual organisations
  - Forming around a given multi-year project
  - Handling large datasets
  - Faced with acquiring and building project infrastructure
- Require infrastructure
  - Larger datasets
  - Data management, data distribution, data processing
  - Challenging a team's ability to produce and maintain infrastructure
- May have access to national and regional infrastructure

**The VO is a paradigm for Supporting interdisciplinary and collaborative research in astronomy and exploiting the full power of growing and emerging data sets**

## The VO is a framework

- For data centers to provide co-operating data services,
- For software providers to offer a variety of compatible analysis and visualization tools and user interfaces

- INAF had more than 90 software packages developed, some public, many “locally” to be engineered
- Raw data is public, but “science ready” data is not yet.
  - Using DOI to suggest share experience and work (software, data, gray articles, .. )

- Sistemi Monolitici non sono la soluzione,  
→ ma neanche il troppo distribuito
- Nessuno ti regala niente  
→ le partnership hanno pro/contro
  
- Domanda e': cosa vuole INAF:
  - Creare una cultura del “Big data analyst” per essere pronti tra 5-10 anni
  - Sviluppare la cultura della Proprieta' Intellettuale su SW e Dati
  - Non Demandare completamente ad “altri”

Grazie per l'attenzione

