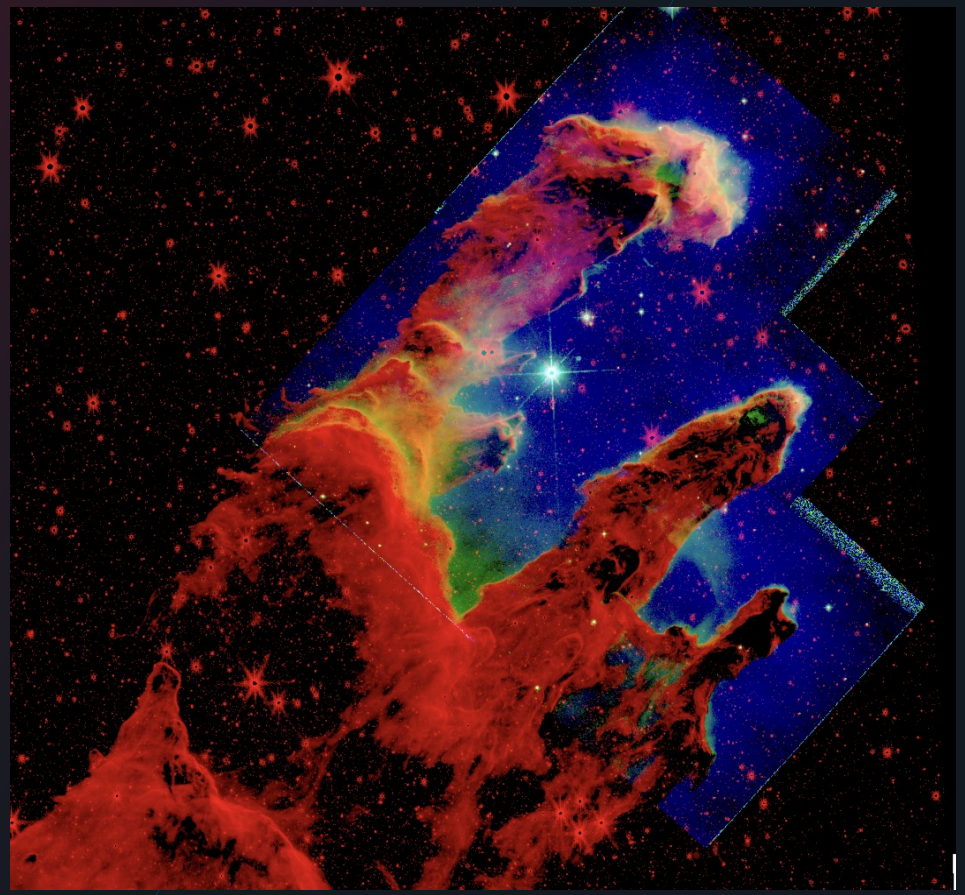
Universe Through Archival Data

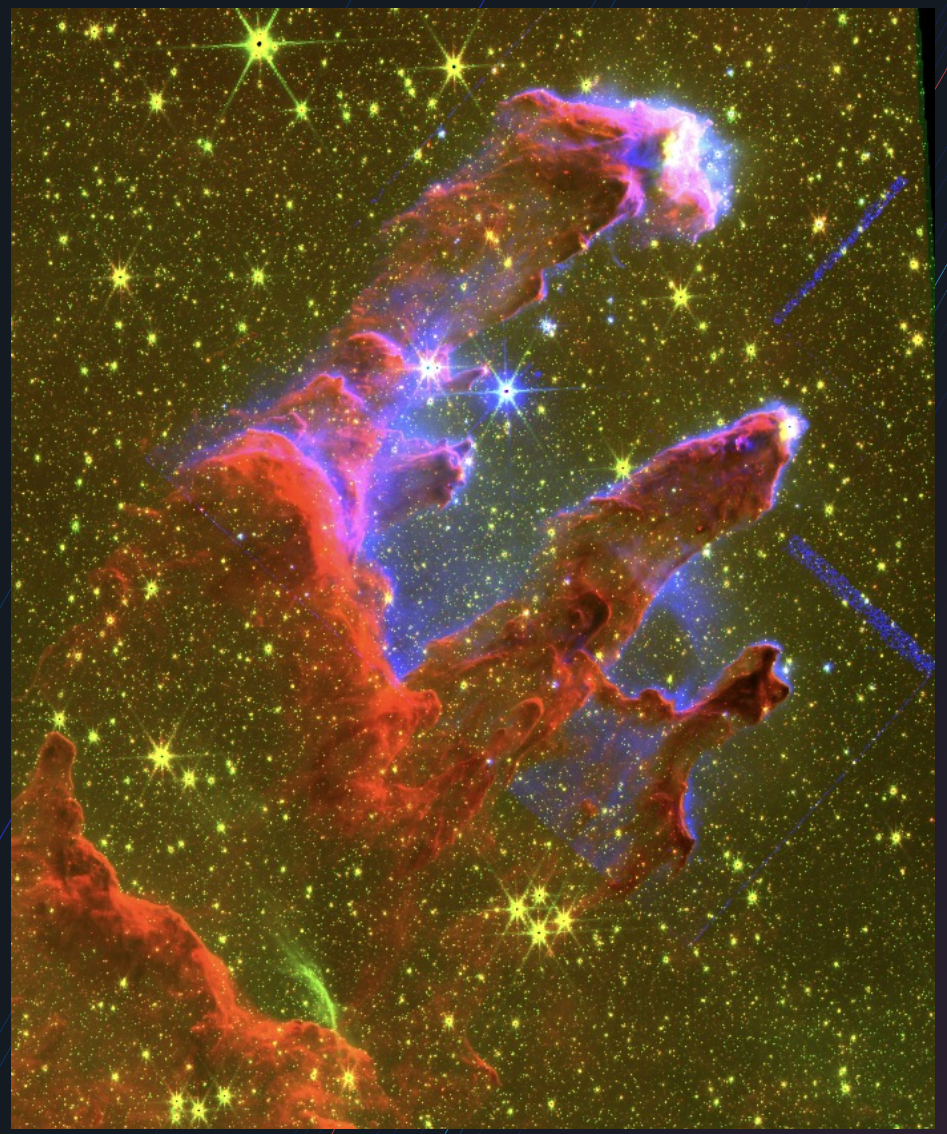
Understanding and Creating Astronomical Images

Pavlo Plotko
2nd ChETEC-INFRA Observational School (ChINOS) 2025
21.07.2025
pavlo.plotko@dzastro.de
plotkopavlo.com

The goal: create an image of M16

Lean about archives and images



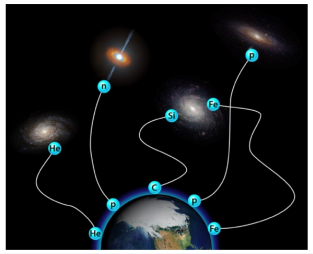


M16 – Pillars of Creation

Who am I?

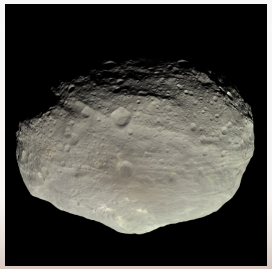
Building and coordinating of all activities with Eastern Europe/Baltic at DZA research (”part-time”)

PhD: theoretical astrophysics; UHECRs; (DESY Germany)



Master's and Bachelor's: Astronomy;
Thesis: Acceleration and non-thermal emission of cosmic rays in SNR (Kyiv, Ukraine)

Before that: Photometric studies of asteroids, stars, and GRB afterglow (Kharkiv, Ukraine)



Organizing:
Astronomy Tournaments
Astrophysics Olympiads
Astrophysics Conferences

Summer schools
Hackathon

Coordinator of the recovery of Ukrainian Astronomy

Web developer and Software developer in the past

Addressing the Elephant in the Room: AI and Chatbots

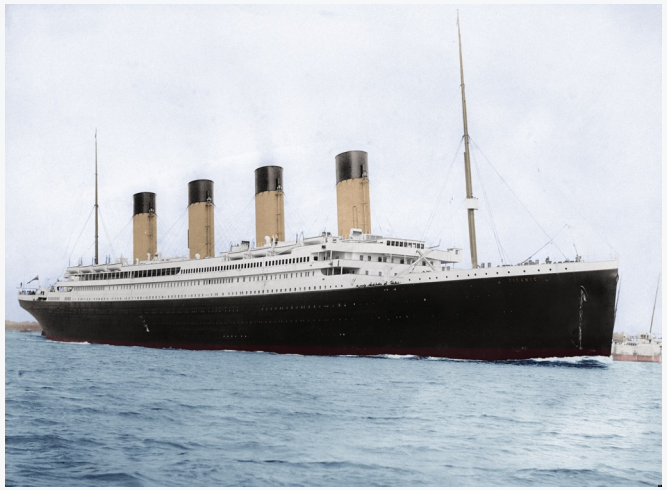
Asking the right questions and understanding the ideas is more important than ever — that's what I aim to do here.

Since you can simply ask ChatGPT (please not the new version of Grok) for step-by-step DS9 instructions — and it works."

https://zapier.com/blog/best-ai-chatbot/

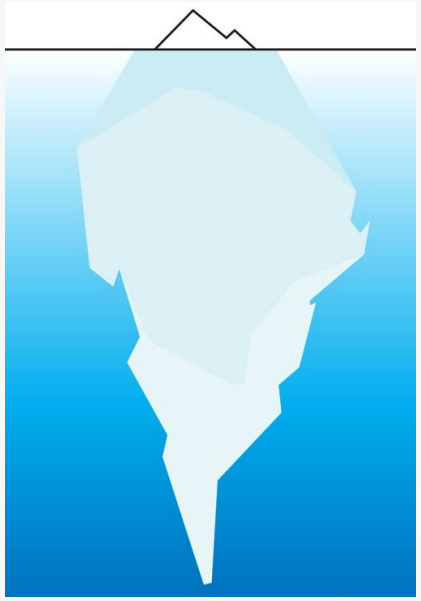
Images and Archives Are Like Icebergs

The beauty is just the tip; the complexity is hidden below.

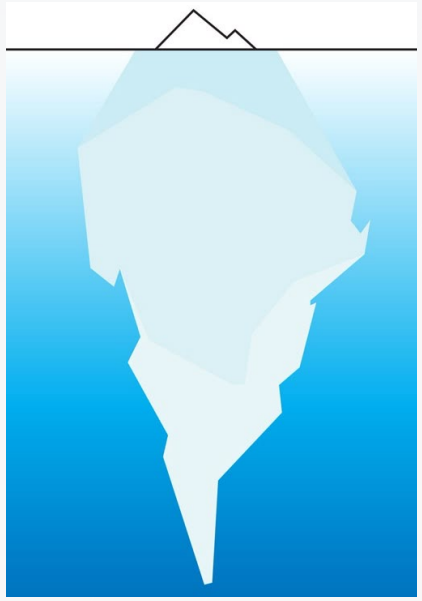


https://commons.wikimedia.org/wiki/File:Titanic_in_color.png

Archive data



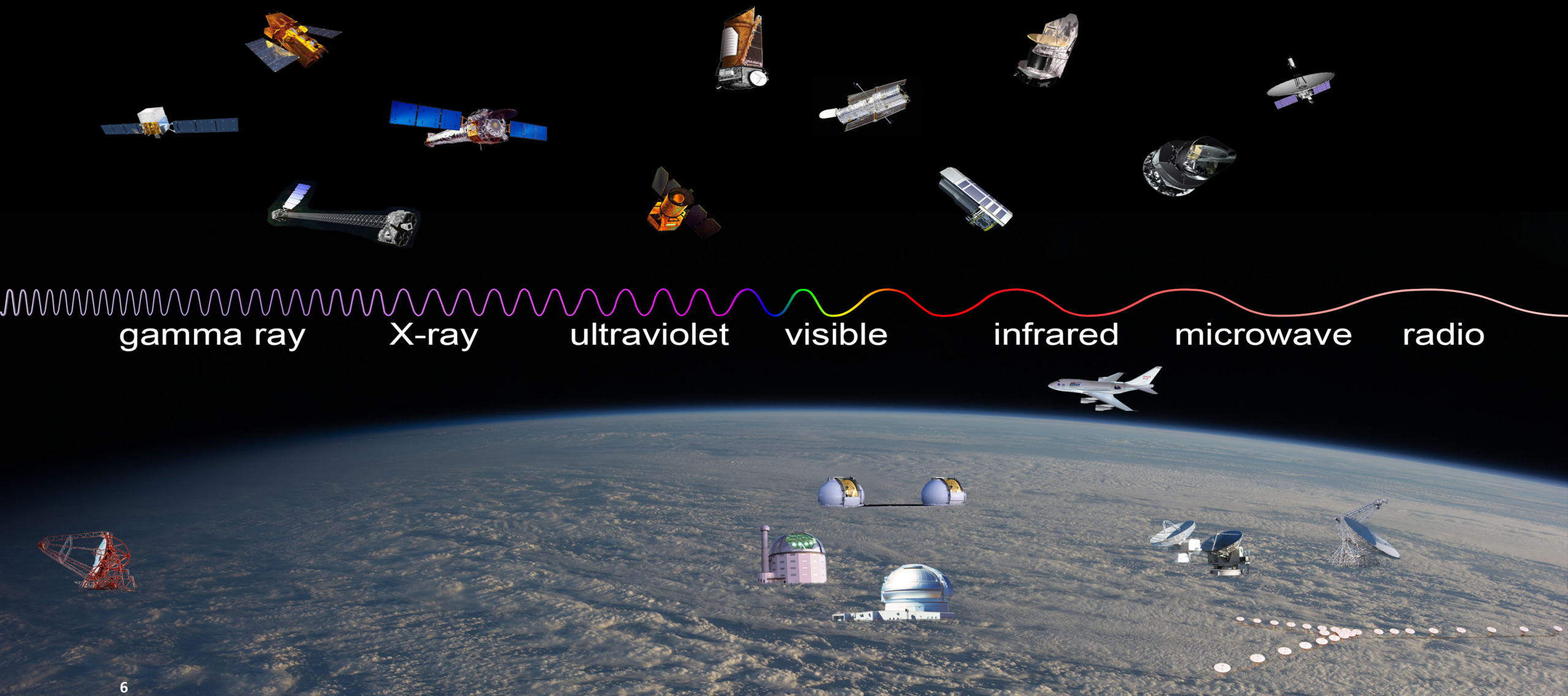
Image



https://www.pinterest.com/pin/tip-of-the-iceberg-poster--189151253084507208/

Huge volumes of astronomical data from multiple missions

https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum_observatories1.html



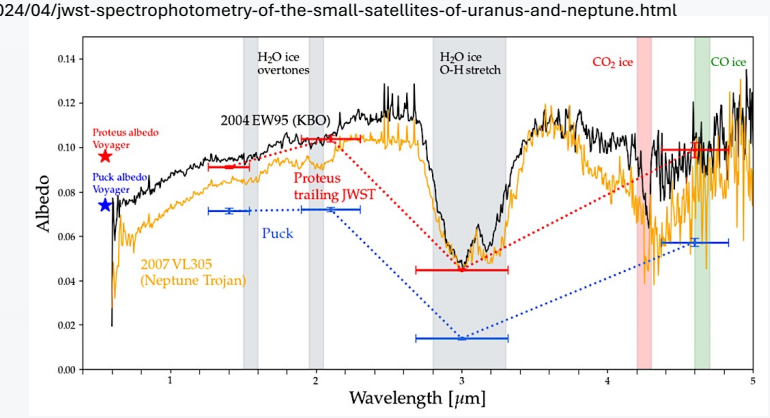
Each ”Detector” has its own archive

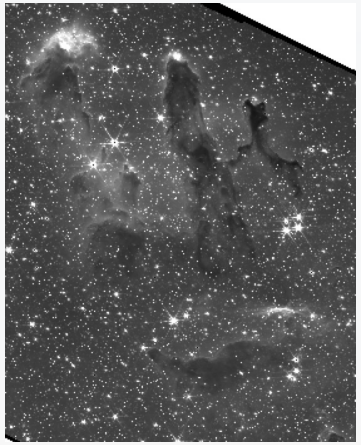
No One-Size-Fits-All

Archives are shaped by:
- Mission goals
- Instruments
- Wavelength coverage

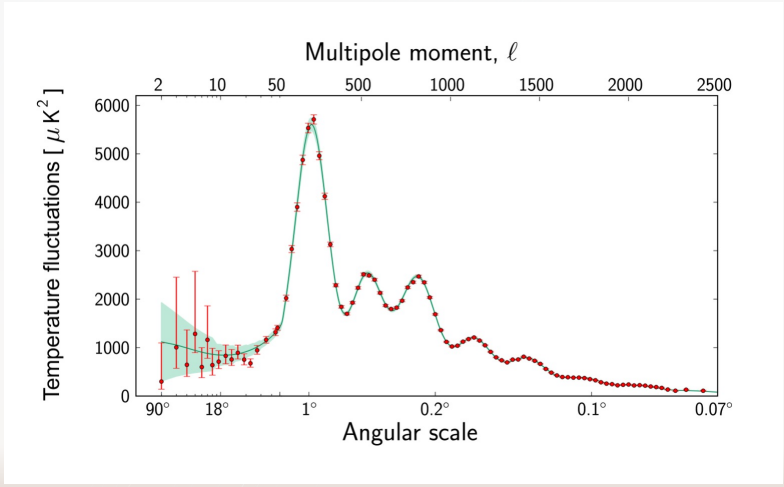
Examples:
Gaia: astrometry & photometry
JWST: IR imaging/spectroscopy
Planck: microwave background
XMM-Newton: X-ray spectra

https://astrobiology.com/2024/04/jwst-spectrophotometry-of-the-small-satellites-of-uranus-and-neptune.html

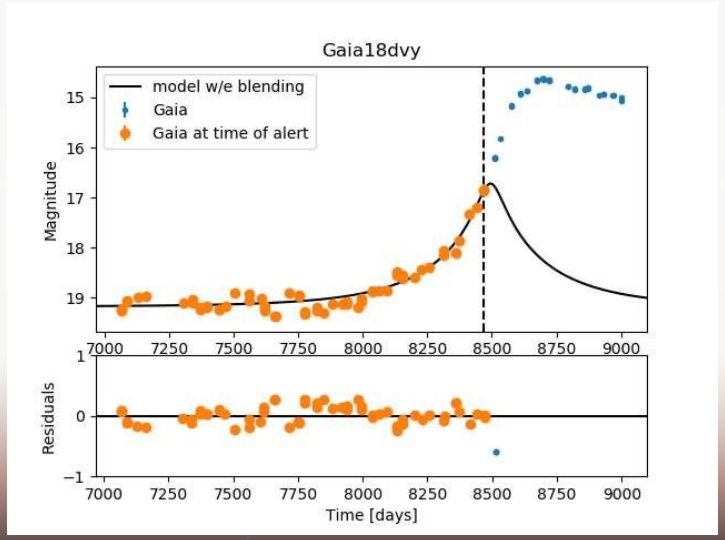




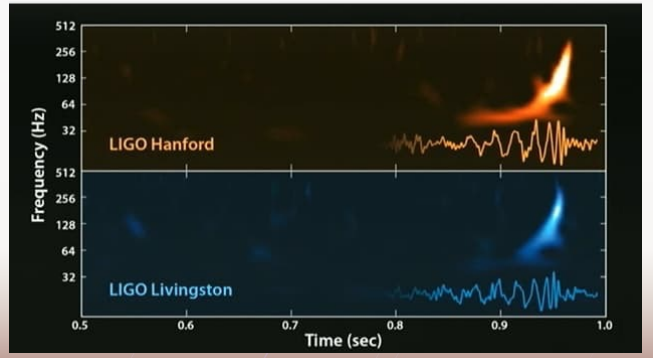
https://sci.esa.int/web/planck/-/51555-planck-power-spectrum-of-temperature-fluctuations-in-the-cosmic-microwave-background



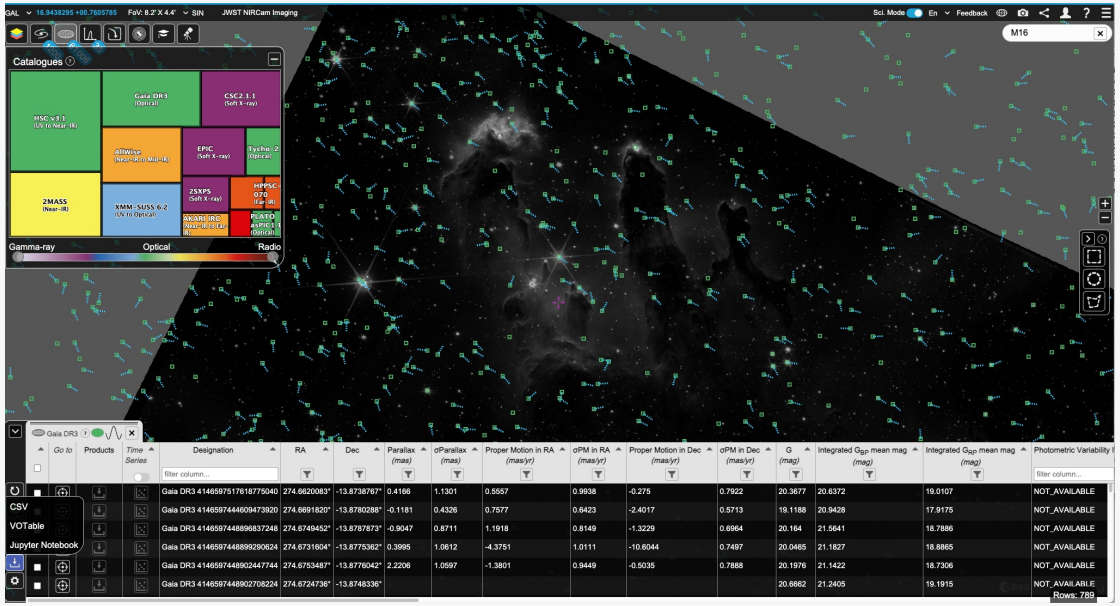
Archives are unique, science-driven, and often incompatible without translation tools



10.48550/arXiv.2201.12209



LIGO/*Phys. Rev. Lett.* **116** 061102



Access Modes

Web interfaces (ESA Sky, MAST, IRSA, etc.)

APIs and programmatic access (e.g., astroquery)

Table downloads, cutout services, cone searches

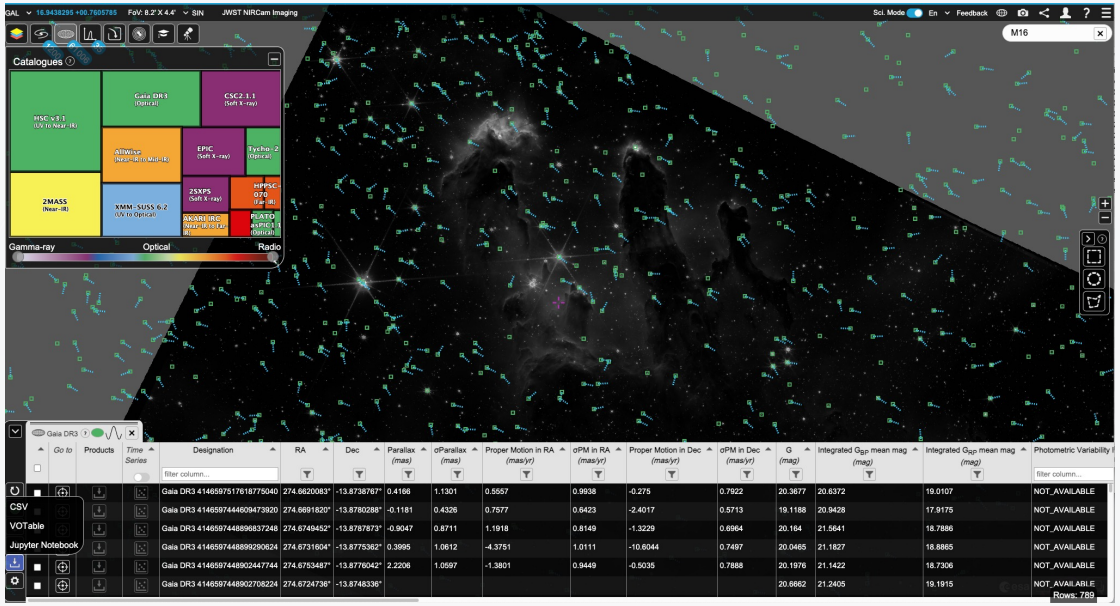
Different missions & archives may support different modes — choose what fits your workflow!

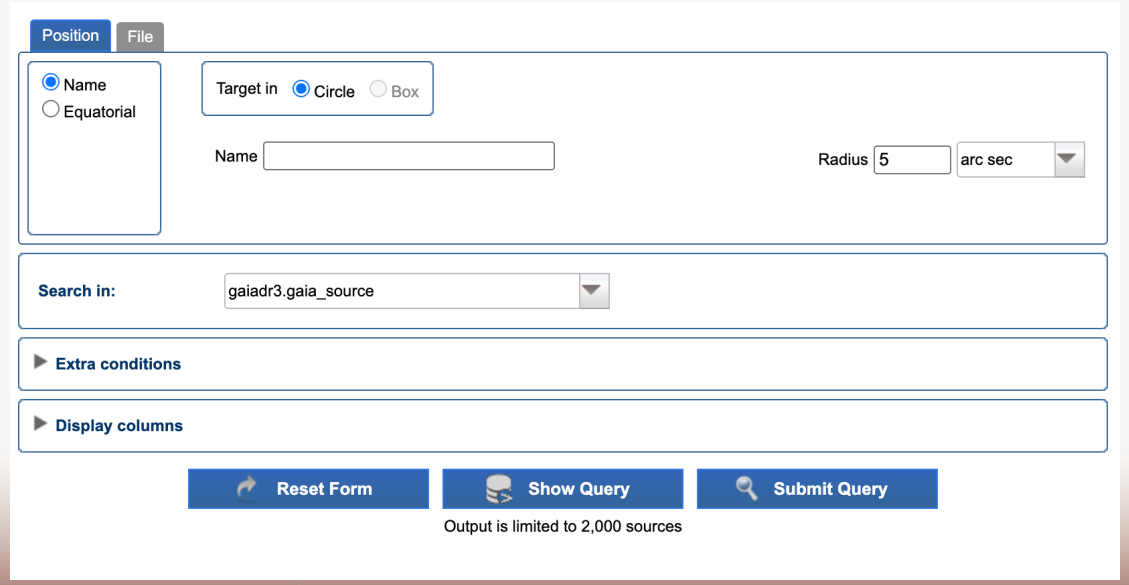
Access Modes

Web interfaces (ESA Sky, MAST, IRSA, etc.)

APIs and programmatic access (e.g., astroquery)

Table downloads, cutout services, cone searches





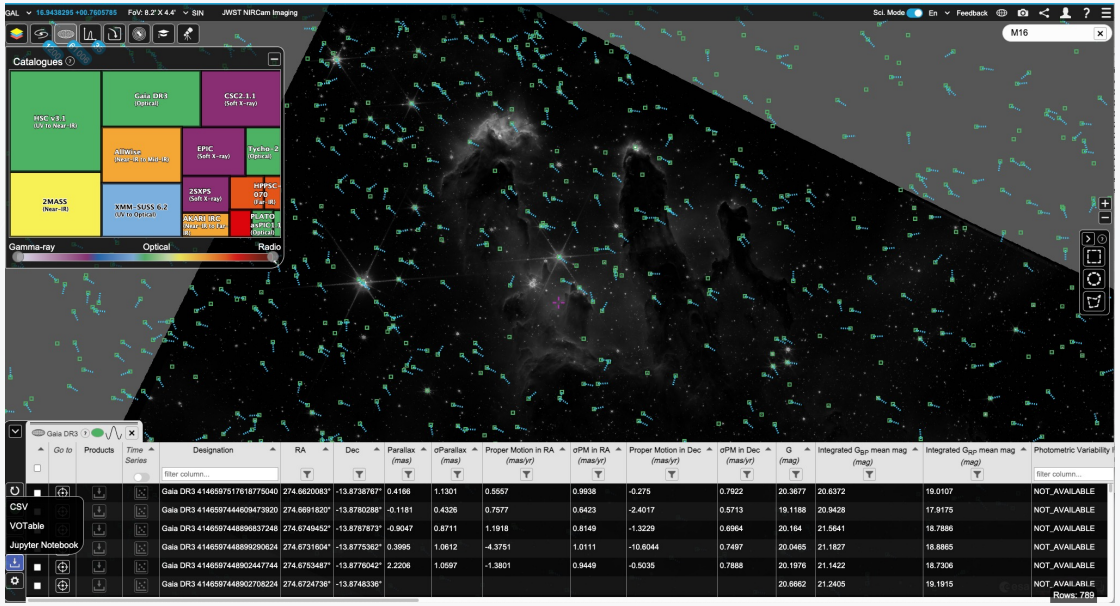
Different missions & archives may support different modes — choose what fits your workflow!

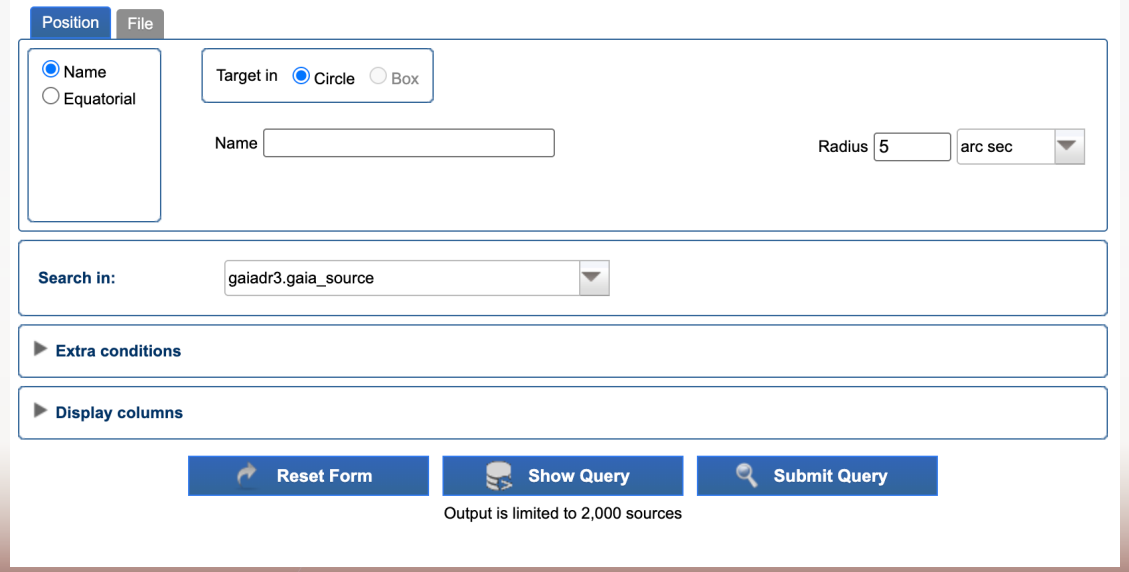
Access Modes

Web interfaces (ESA Sky, MAST, IRSA, etc.)

APIs and programmatic access (e.g., astroquery)

Table downloads, cutout services, cone searches





Different missions & archives may support different modes — choose what fits your workflow!

```
In [7]: queryGJ473 = """SELECT source_id, phot_g_mean_mag, parallax, pmra, pmdec, ruwe
FROM gaiadr3.gaia_source
WHERE
CONTAINS(
        POINT('ICRS',ra,dec),
        CIRCLE(
                'ICRS',
                COORD1(EPOCH_PROP_POS(188.3223395,9.0210534,227.9000,-1730.0000,203.0000,1.0000,2000,2016.0)),
                COORD2(EPOCH_PROP_POS(188.3223395,9.0210534,227.9000,-1730.0000,203.0000,1.0000,2000,2016.0)),
                5./3600.)
)=1"""

In [8]: job = Gaia.launch_job(queryGJ473)
r = job.get_results()
r
```

No One-Size-Fits-All

Archives are shaped by:

- Mission goals
- Instruments
- Wavelength coverage

Examples:

Gaia: astrometry & photometry
JWST: IR imaging/spectroscopy
Planck: microwave background
XMM-Newton: X-ray spectra

Access Modes

Web interfaces (ESA Sky, MAST, IRSA, etc.)

APIs and programmatic access (e.g., astroquery)

Table downloads, cutout services, cone searches

Open vs Restricted Archives

Open archives: Freely accessible data (e.g., Hubble, Gaia, JWST)

Restricted (proprietary): Data under embargo for 6–12 months (or years…)

Key point: Always check data rights before using or publishing.

Developing and Maintaining

Who Should Build and Maintain Archives?

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

🏛️ **Dedicated institutions**

- ✅ Long-term expertise & infrastructure
- ❌ Needs sustained funding

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

🏛️ **Dedicated institutions**

- ✅ Long-term expertise & infrastructure
- ❌ Needs sustained funding

🤝 **Community/Virtual Observatory**

- ✅ Shared standards & effort
- ❌ Harder to coordinate

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

🏛️ **Dedicated institutions**

- ✅ Long-term expertise & infrastructure
- ❌ Needs sustained funding

🤝 **Community/Virtual Observatory**

- ✅ Shared standards & effort
- ❌ Harder to coordinate

How Long Should Data Stay Online?

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

🏛️ **Dedicated institutions**

- ✅ Long-term expertise & infrastructure
- ❌ Needs sustained funding

🤝 **Community/Virtual Observatory**

- ✅ Shared standards & effort
- ❌ Harder to coordinate

How Long Should Data Stay Online?

📈 **Why keep data?**

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 Mission team

✅ Knows instruments & data

❌ Limited by mission budget & timeline

🏛️ Dedicated institutions

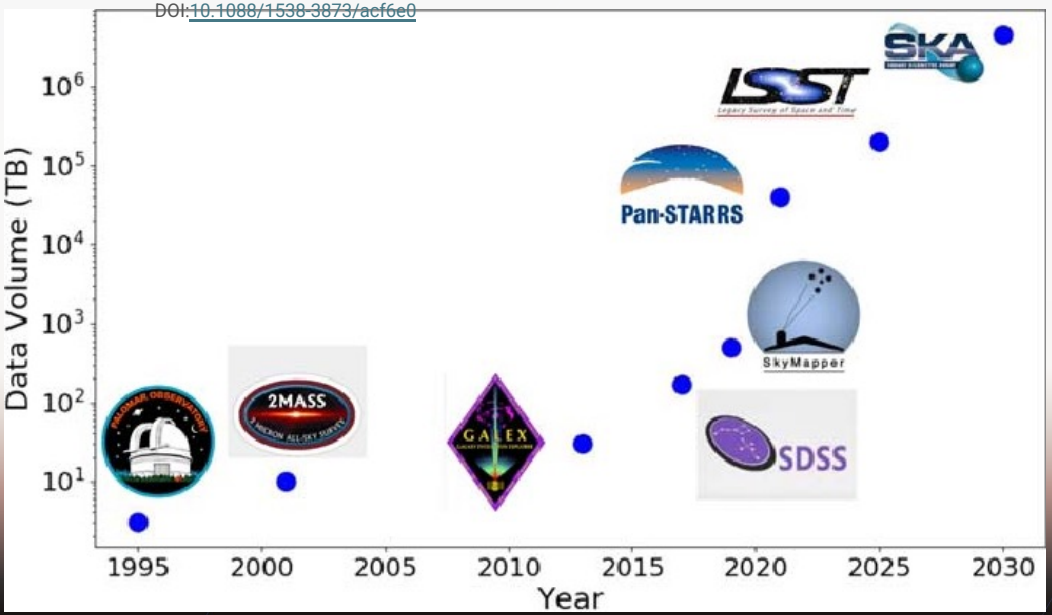
✅ Long-term expertise & infrastructure

❌ Needs sustained funding

🤝 Community/Virtual Observatory

✅ Shared standards & effort

❌ Harder to coordinate



How Long Should Data Stay Online?

📈 Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

📉 Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Who Should Build and Maintain Archives?

🚀 **Mission team**

- ✅ Knows instruments & data
- ❌ Limited by mission budget & timeline

🏛️ **Dedicated institutions**

- ✅ Long-term expertise & infrastructure
- ❌ Needs sustained funding

🤝 **Community/Virtual Observatory**

- ✅ Shared standards & effort
- ❌ Harder to coordinate

DOI:10.1088/1538-3873/acf6e0

How Long Should Data Stay Online?

📈 **Why keep data?**

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

📉 **Why not?**

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

If we had unlimited money...

How Long Should Data Stay Online?

Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

If we had unlimited money...

Could we store all the data?

How Long Should Data Stay Online?

Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

If we had unlimited money...

Could we store all the data?
Yes.

How Long Should Data Stay Online?

Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

If we had unlimited money...

Could we store all the data?
Yes.

But should we?

How Long Should Data Stay Online?

Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

If we had unlimited money...

Could we store all the data?
Yes.

But should we?
Probably not.

How Long Should Data Stay Online?

Why keep data?

- Legacy science, reproducibility, cross-mission studies
- New discoveries from old data

Why not?

Storage & maintaining cost money

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Making everything more complex: Carbon footprint (ALMA example)

Storage of 1TB of data -> 2000kg/yr of CO2

Transfer of 1 GB of data -> 3kg

@Wiki

https://www.mur.gov.it/sites/default/files/2024-03/Astronomical_data_SIGRE.pdf

Making everything more complex: Carbon footprint (ALMA example)

Storage of 1TB of data -> 2000kg/yr of CO2

Transfer of 1 GB of data -> 3kg

A 100 GB dataset in ALMA Archive in 5 yr generates 13000kg of CO2
corresponding to CO_2 generated by 3 cars driven continuously per 1yr
CO_2 absorbed by 215 trees in 10 yr

@Wiki

https://www.mur.gov.it/sites/default/files/2024-03/Astronomical_data_SIGRE.pdf

Making everything more complex: Carbon footprint (ALMA example)

Storage of 1TB of data -> 2000kg/yr of CO2

Transfer of 1 GB of data -> 3kg

A 100 GB dataset in ALMA Archive in 5 yr generates 13000kg of CO2
corresponding to CO_2 generated by 3 cars driven continuously per 1yr
CO_2 absorbed by 215 trees in 10 yr

the ALMA archive we have more than 60000 datasets (10GB-1TB size) =
= CO2 absorbed by trees covering the whole area of Florence for 10yr

@Wiki

https://www.mur.gov.it/sites/default/files/2024-03/Astronomical_data_SIGRE.pdf

No One-Size-Fits-All

Archives are shaped by:

- Mission goals
- Instruments
- Wavelength coverage

Examples:
Gaia: astrometry & photometry
JWST: IR imaging/spectroscopy
Planck: microwave background
XMM-Newton: X-ray spectra

Access Modes

Web interfaces (ESA Sky, MAST, IRSA, etc.)

APIs and programmatic access (e.g., astroquery)

Table downloads, cutout services, cone searches

Open vs Restricted Archives

Open archives: Freely accessible data (e.g., Hubble, Gaia, JWST)

Restricted (proprietary): Data under embargo for 6–12 months (or years…)

Key point: Always check data rights before using or publishing.

Developing and Maintaining

- Who should build and maintain archives?

  mission team; Dedicated institutions; Community-driven or VO (Virtual Observatory)

- How long should data stay online?

Centre de Données astronomiques de Strasbourg (CDS)

Founded in 1972, located in Strasbourg, France
A multidisciplinary team of astronomers, software engineers, and data librarians
Curates and distributes astronomical data from publications, catalogs, and reference images
Provides three major services:

SIMBAD – object database with cross-identifiers and bibliography (~7.5 million objects)
VizieR – access to catalogs, observation logs, tables, images, spectra, time series
Aladin – interactive sky atlas and Virtual Observatory (VO) portal
Xmatch – a new tool allowing astronomers to efficiently cross-identify sources between overlapping catalogues

ESA Sky Interface

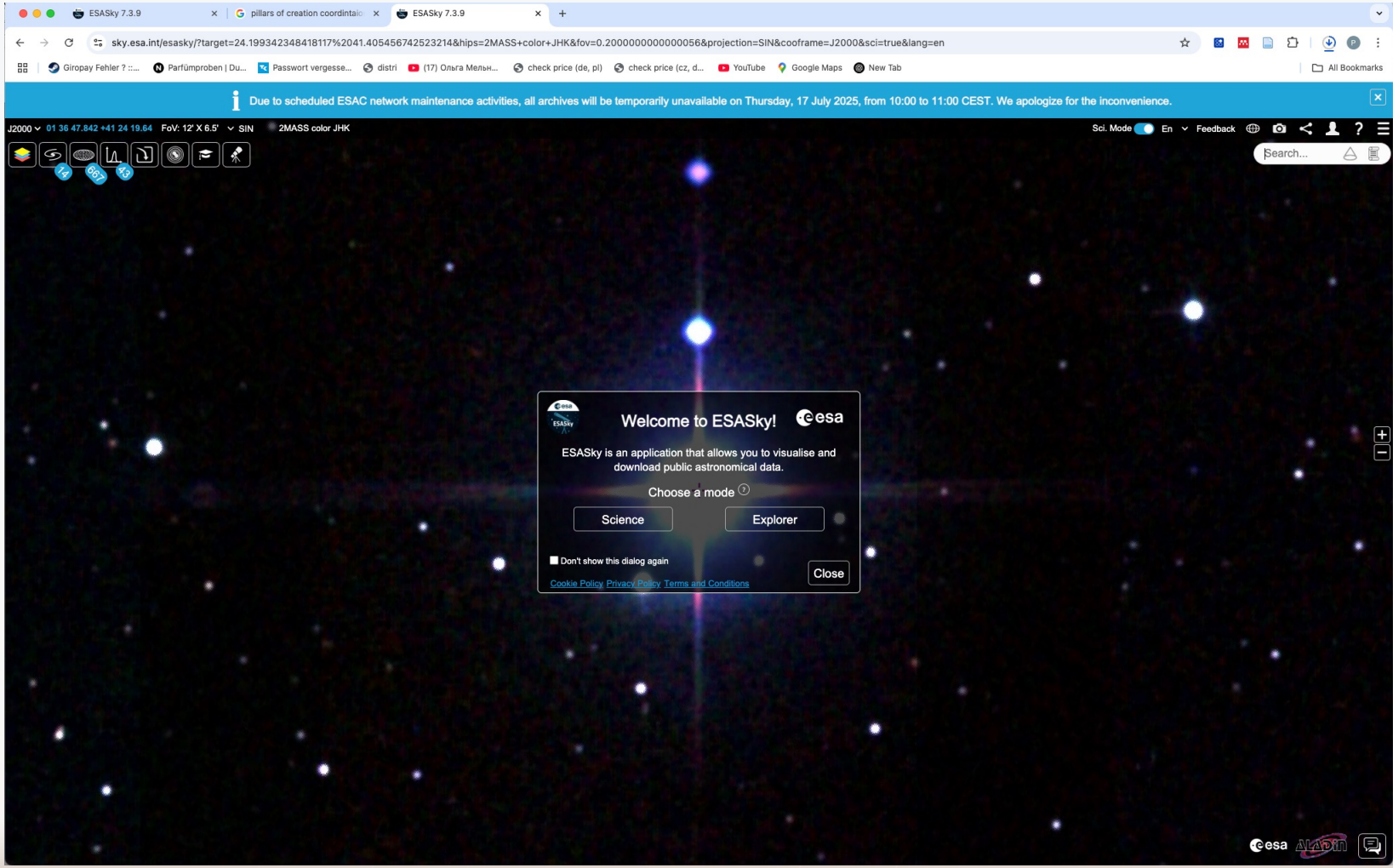
Unified data access portal for ESA missions (HST, JWST, Gaia, XMM-Newton, Planck, etc.)
Allows sky exploration, image previews, and FITS downloads
Combines professional-grade tools with an accessible, visual interface

Focusing on ESA SKY

Why ESA? Unified interface, high usability, access to major missions
Overview of https://sky.esa.int
Supports Hubble, JWST, Gaia, XMM-Newton, Planck, and more

URL: https://sky.esa.int

Main features:
Sky navigation
Object search
Filter selection
FITS download
Multi-mission layering



Websites with Professional FITS Image Archives

Professional astronomers use these websites to store and distribute their data, including FITS images. Many of these websites expect people to search for images of specific objects; a list of example objects is presented after this list of websites.

ESASky http://sky.esa.int/

GALEX View (requires Adobe flash)
https://galex.stsci.edu/GalexView/

Herschel Database in Marseille http://hedam.lam.fr/

Herschel User Provided Data Products
http://www.cosmos.esa.int/web/herschel/user-provided-data-products

Mikulski Archive for Space Telescopes (MAST) Portal
https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html

Sloan Digitized Sky Survey https://dr12.sdss.org/fields

NASA/IPAC Infrared Science Archive
http://irsa.ipac.caltech.edu/images.html

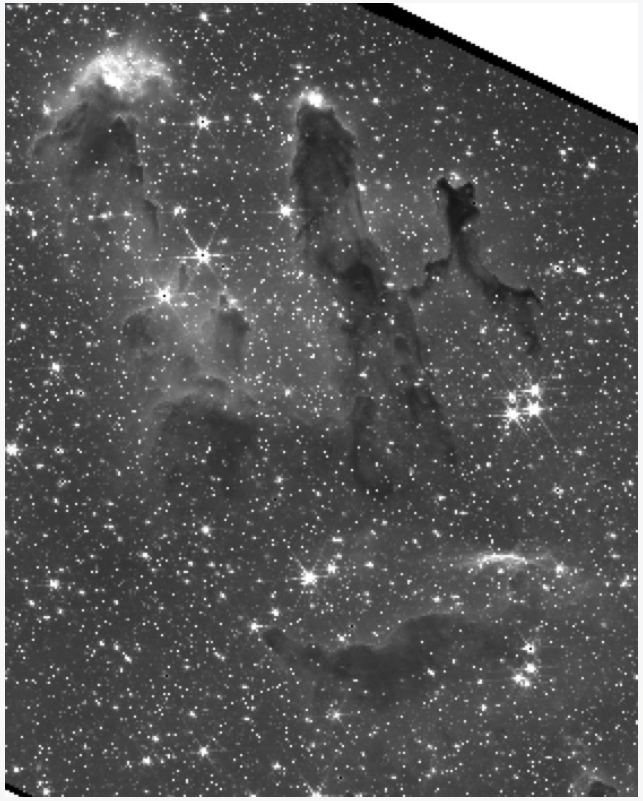
NASA/IPAC Extragalactic Database
http://ned.ipac.caltech.edu/

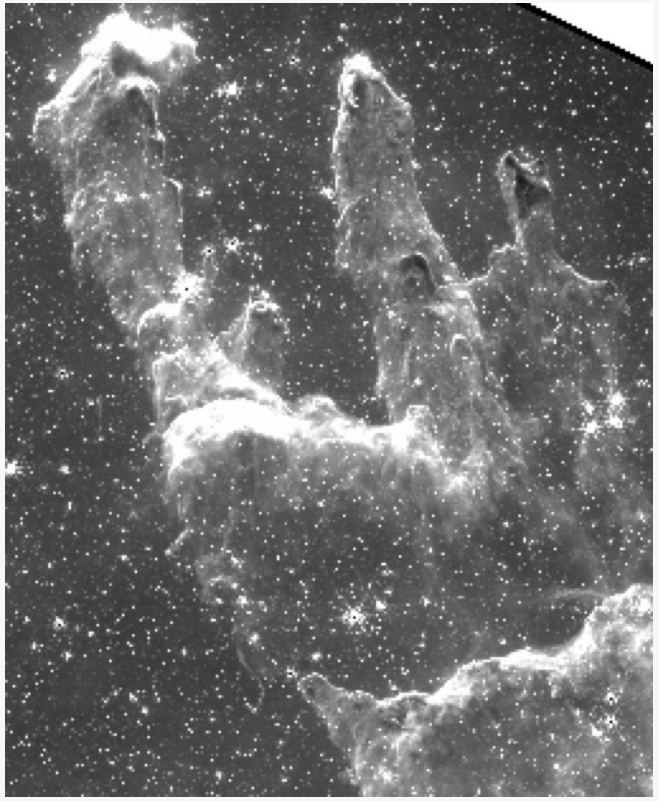
PanSTARRS-1 Image Access
https://ps1images.stsci.edu/cgi-bin/ps1cutouts

Spitzer 24-160 Micron Data for the Herschel-SPIRE Local Galaxies Guaranteed Time Programs
http://www.jb.man.ac.uk/~gbendo/exchange/SpitzerData/spitzerdata_main.html

UKIRT InfraRed Deep Sky Surveys
http://wsa.roe.ac.uk:8080/wsa/getImage_form.jsp

Use JWST and Hubble data of the Pillars of Creation.





How do we go from grayscale to color?

What is an image?*

* Raster

What is an image?

https://www.brandonrohrer.com/images_to_numbers

What is an image?

https://www.brandonrohrer.com/images_to_numbers

What is an image?

https://www.brandonrohrer.com/images_to_numbers

What is an image?

https://www.brandonrohrer.com/images_to_numbers

What is an image?

https://www.brandonrohrer.com/images_to_numbers

Color system

Pillars of Creation (NIRCam Image)

VIEW ALL IMAGES >

Zoom Image

Download Options

Full Res (For Print), 8423 X 14589, TIF (163.43 MB)

Full Res (For Display), 8423 X 14589, PNG (152.33 MB)

1155 X 2000, PNG (4.73 MB)

Why “For Display” and “For Print”?

Color system

What is the difference?

Color system

What is the difference?

Screens; Additive

Printing; Subtractive

https://www.epackprinting.com/cmyk-subtractive-color-mixing/

Color system

What is the difference?

Screens; Additive

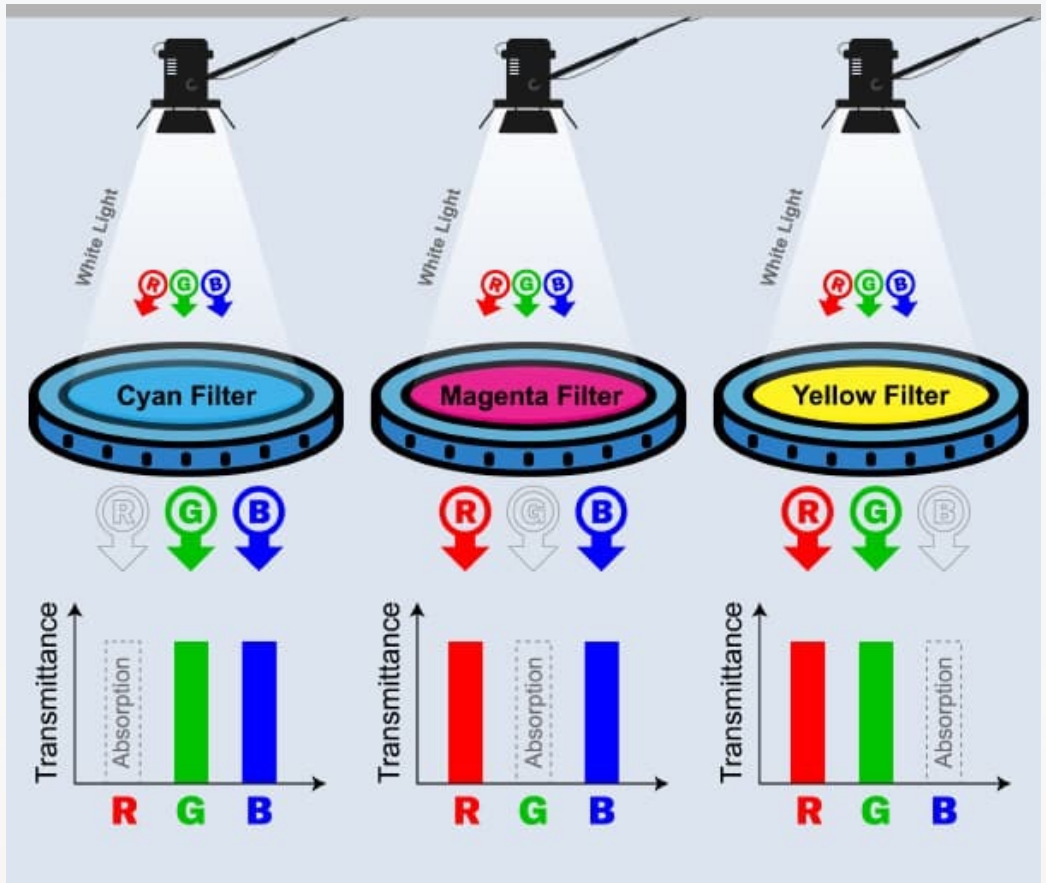
Printing; Subtractive

https://www.epackprinting.com/cmyk-subtractive-color-mixing/

Color system

What is the difference?

Screens; Additive



Printing; Subtractive

https://www.epackprinting.com/cmyk-subtractive-color-mixing/