



# EXPLORE

Innovative Scientific Data Exploration and  
Exploitation Applications for Space Sciences

Nick Cox @SCIOPS 2022



This project has received funding from the European Union's Horizon 2020 research and innovation programme under agreement No 101004214



# SPACE-30-SCI-2020: Scientific data exploitation

## **Specific Challenge:**

Support the data exploitation of European missions and instruments, in conjunction, when relevant, with international missions.

## **Expected Impact:**

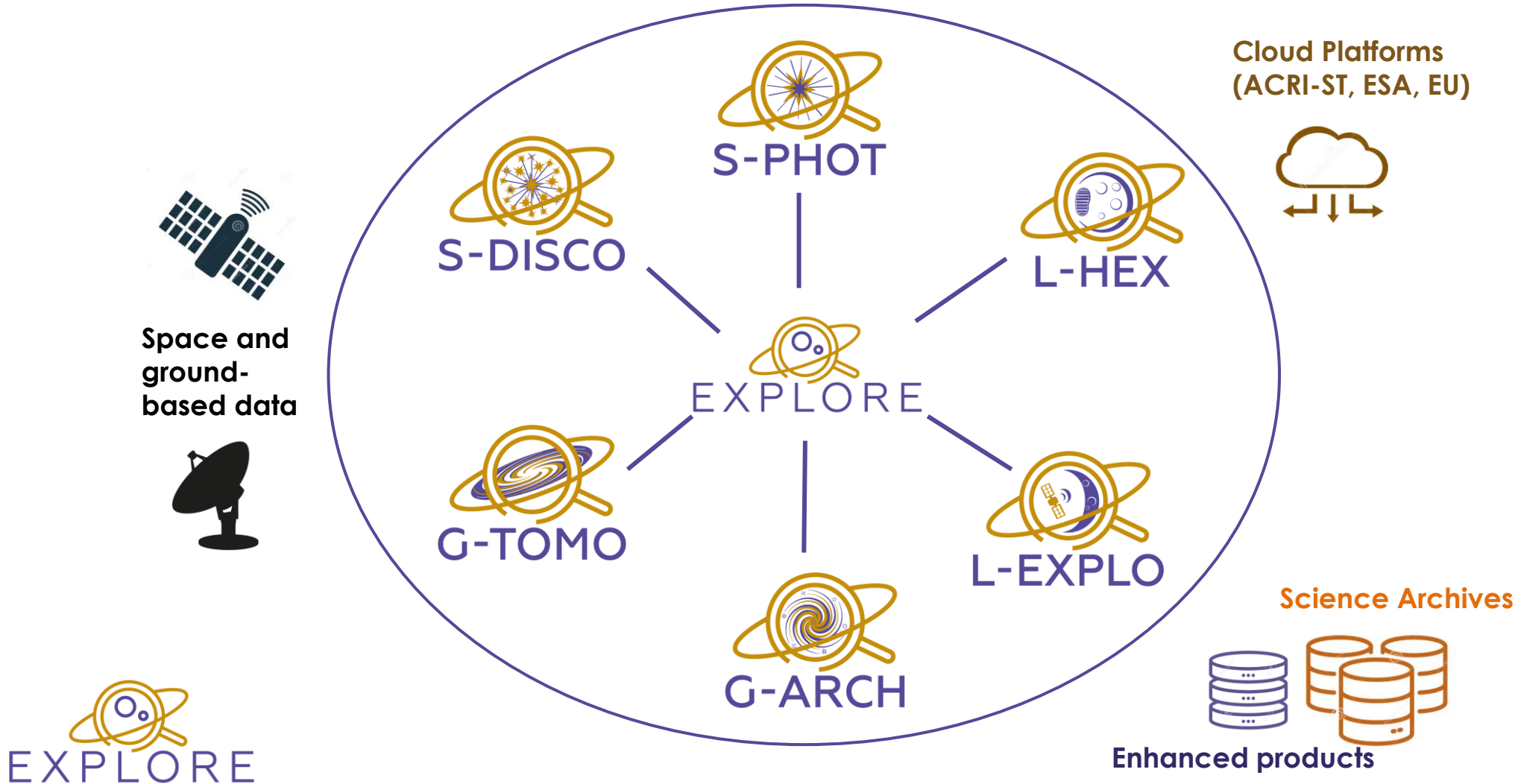
A higher number of scientific publications based on Europe's space data, high-level data products made available through appropriate archives, and tools and methods developed for the advanced processing of data. Proposals are also expected to add value to existing activities on European and international levels, and to enhance and broaden research partnerships.

# SPACE-30-SCI-2020 → EXPLORE

“EXPLORE’s main objective is to develop and deploy a suite of scientific applications [...], to achieve efficient [...] exploitation of scientific data from astrophysics and planetary space missions, [...].”

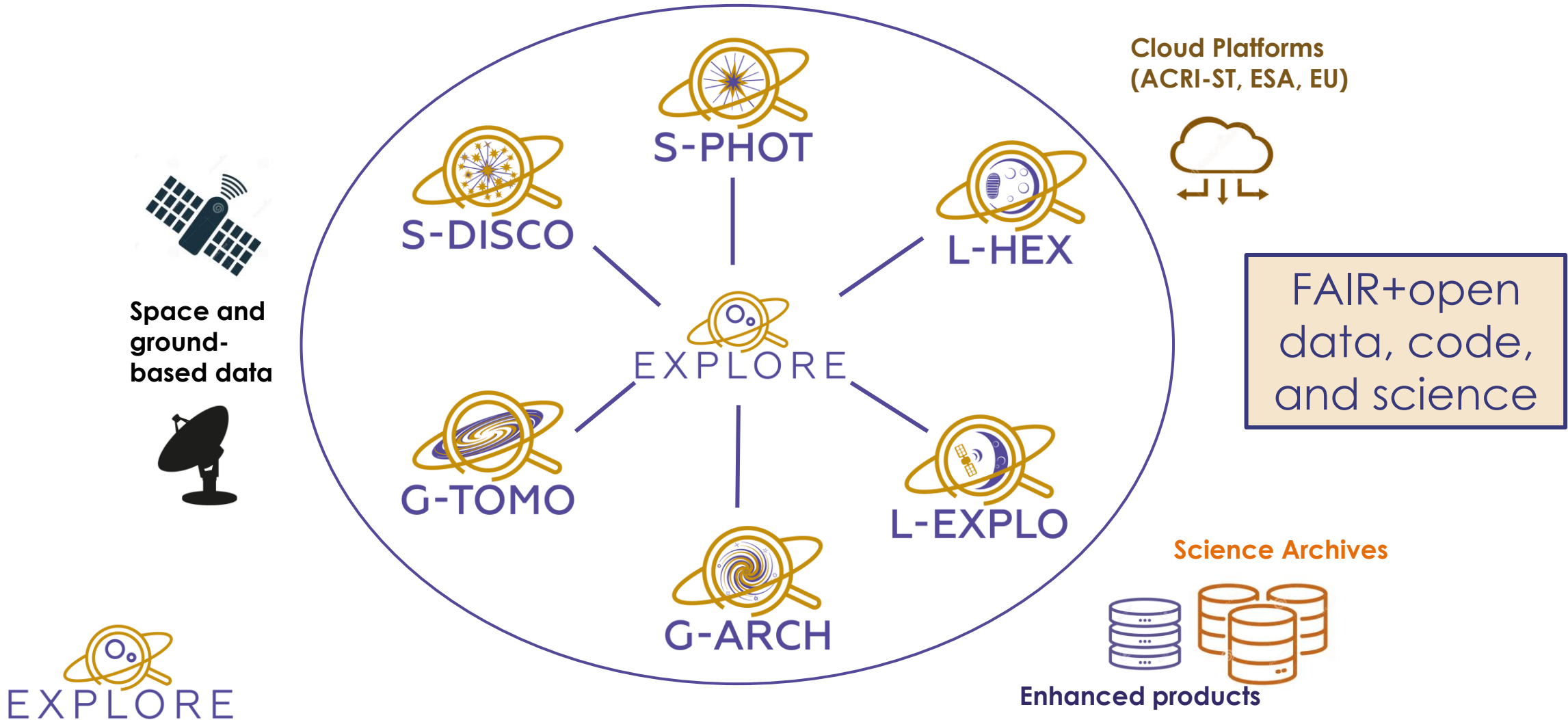
# EXPLORE – Space Science in the Cloud

Lunar exploration and Gaia science applications powered with advanced **visualization** and **machine learning** features



# EXPLORE – Space Science in the Cloud

Lunar exploration and Gaia science applications powered with advanced **visualization** and **machine learning** features



# Science Platforms & EXPLORE-platform

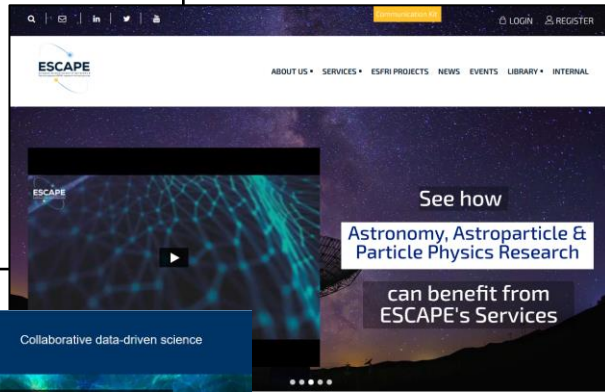
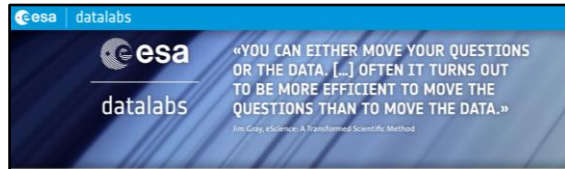
**Science Platforms** → scientific applications in the 'cloud' (close to the data)

For **space sciences**

- US: SciServer, CyVerse (generic)
- Europe: ESA Datalabs, ESCAPE SAP  
→ EOSC ecosystem

EXPLORE **dev/test** platform for SDA (with limited resources 🧑🧑🧑 & 🖥️)

Beyond project EXPLORE platform could offer **bespoke niche services**



# EXPLORE dev platform

## Space browser

## Applications

The screenshot shows the SDAS (Software Data Access System) interface for the G-Tomo Mini application. The title "SDAS" is at the top right. The application name "G-Tomo Mini" is prominently displayed. Below it, there are sections for "Description" and "operations". The "Description" section includes fields for "G-Tomo SDA", "HAS OUTPUT" (with a checkbox), "Creator" (EXPLORE), and "Licence" (Apache 2.0). The "operations" section has "OPEN" and "STOP" buttons. A "METADATA" link is also visible.

The screenshot shows the Space browser interface. On the right is a grayscale map of the Moon with various lunar features labeled, such as MARE SERENITATIS, MARE CRISLUM, MARE TRANQUILLITATIS, and MARE AUSTRALE. A blue box highlights a specific area on the map. On the left is a sidebar with a search bar and a list of products. The selected product is "Clementine-LWIR". Below the search bar, there are sections for "Required" and "Optional" parameters, including "PRODUCT TYPE" (RDR), "MIN EMISSION ANGLE", "MIN INCIDENCE ANGLE", "MIN PHASE ANGLE", "MAX OBSERVATION TIME", and "MIN OBSERVATION TIME". A "Q SEARCH" button is at the bottom.

## User workspace

The screenshot shows the "My Files" user workspace interface. It features a header "My Files" and a subtitle "Navigate through all files you have access to in the Explore platform's". Below this, there is a "My user space" section with a "PERMISSIONS" tab. It includes buttons for "upload file" and "add folder". A list of folders is shown, each with a name, a lock icon, and a delete icon. The folders are: "TestFiles", "app\_data", "g-tomo\_20211206T143629", "g-tomo\_20211206T150535", "g-tomo\_20211206T155223", "g-tomo\_mini\_20211206T160245", "g-tomo\_mini\_20211206T160252", "g-tomo\_mini\_20211206T160406", "g-tomo\_mini\_20211206T160530", "g-tomo\_mini\_20211206T160534", "g-tomo\_mini\_20211206T161405", and "g-tomo\_mini\_20211206T162802".

The screenshot shows the "SELECT A SOLAR BODY" interface. It features two circular images: the Moon and Mars. Below the images, the word "MOON" is displayed, indicating the selected body.

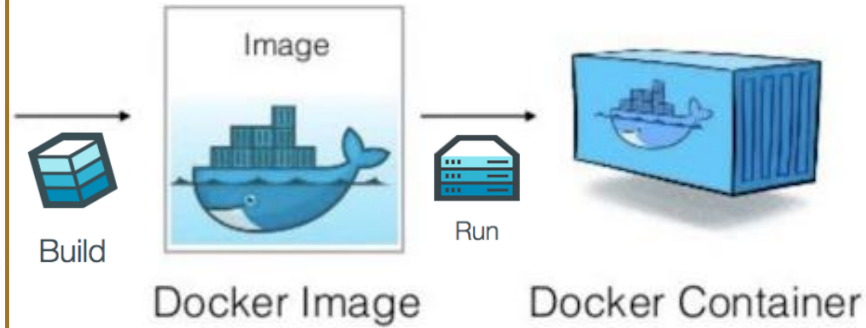
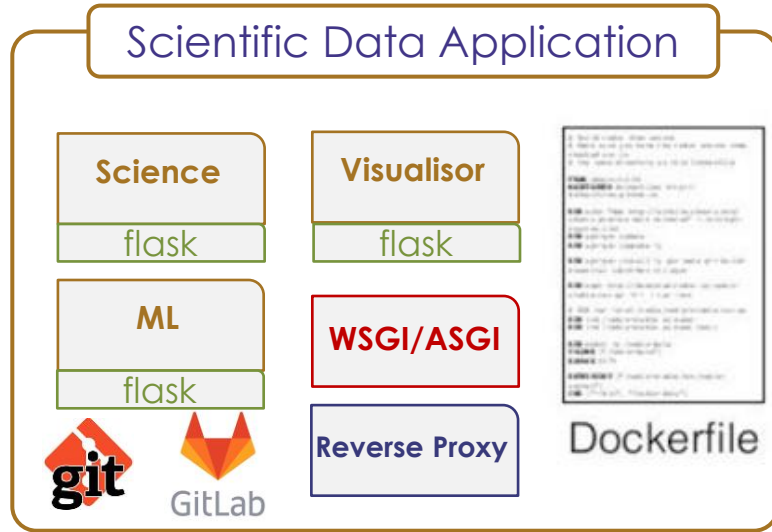
<https://explore-platform.eu>

# Scientific Data Applications (SDAs)?

- Applications / tools for scientific data exploration, visualisation, analysis
- Deployed on science 'cloud' platforms and accessed through a web frontend interface (UI) or API.
- Let users interact remotely with data (bringing the users to the tools and data rather than bringing the tools/data to the user)
- Container approach to create, deploy, and share open-source, interoperable SDAs
- SDAs to demonstrate and promote, leading by example, space science exploitation and uptake of science platforms.



# dev-build-deploy process



[explore-platform.eu](https://explore-platform.eu)  
[datalabs.esa.int](https://datalabs.esa.int)  
...



# Scientific Data Applications

Led by Science Experts



**G-Arch:** Galactic Archaeology



**G-Tomo:** Galactic Interstellar Tomography



**S-Phot:** Stars and their Blue/Red Excess



**S-Disco:** Spectral Discovery for Stars



**L-Explo:** Exploring the Moon with multi-scale data



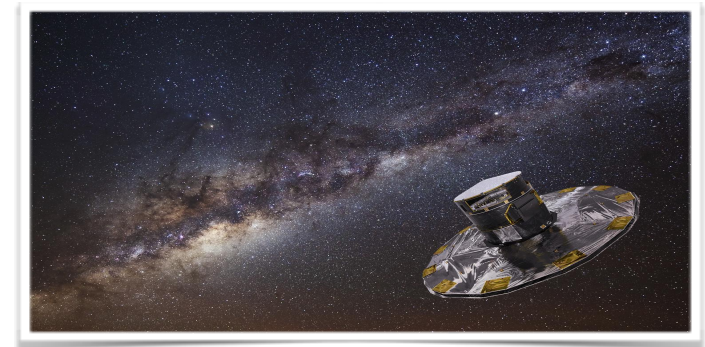
**L-Hex:** Lunar Human Exploration tools





# EXPLORE SDA – S-Disco

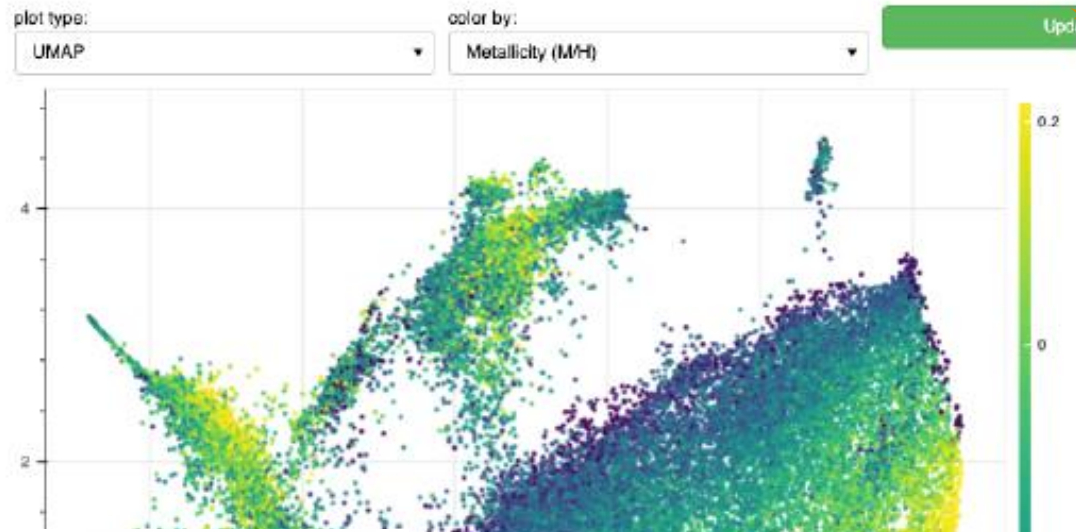
- Surveys are gathering complex multidimensional data for millions/billions of sources.
  - One can no longer examine them all by eye.
  - ML/AI can be used to find “the ones that do not look like the others” = anomalies/outliers/novelties.
  - The purpose is to promote a new way to look at data, and create a channel for the discovery of new phenomena.
- 
- GAIA and APOGEE data retrieved for all overlapping sources.
  - algorithm hyper-parameters explored and selected.



# EXPLORE SDA – S-Disco



- Total W score - The mean (feature) pairwise distance
- Spectral NN W score - The mean (feature) pairwise distance from 250 (feature) NN
- Spatial NN W score - The mean (feature) pairwise distance from 250 (Euclidean) NN



plot type:

- UMAP
- Galactic plane
- Galactic side view
- HR diagram

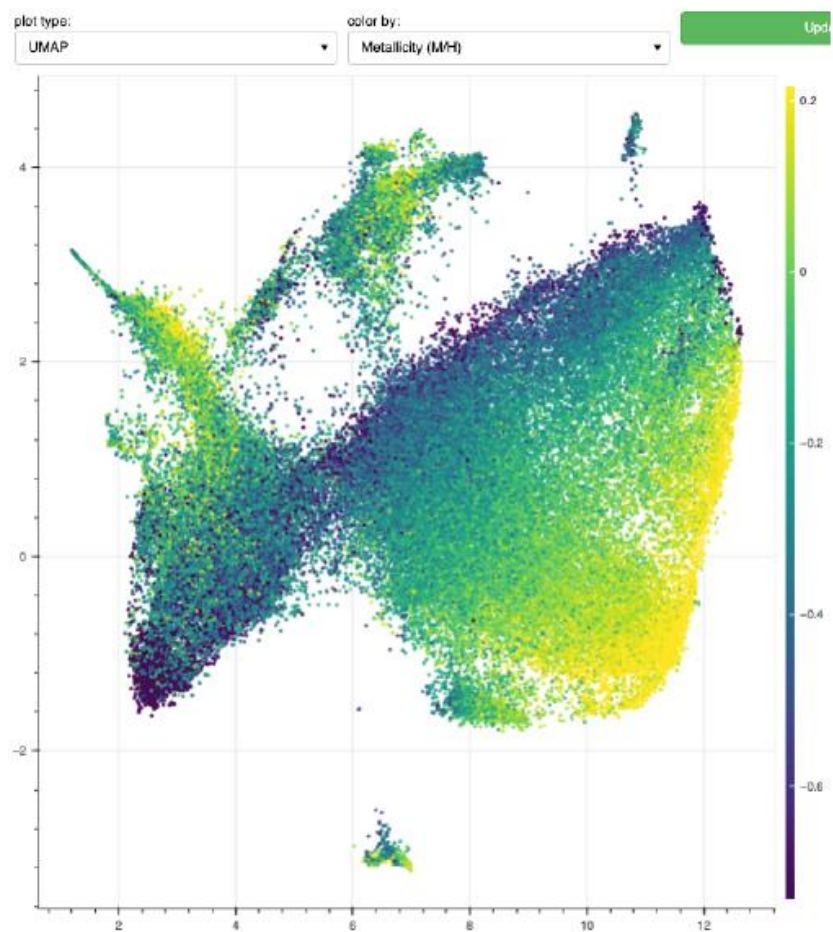
- total weirdness score
- 250 nn spectral weirdness score
- 250 nn spatial weirdness score
- Magnitudes
- effective nu
- BP-RP color
- Distance
- fit MSE
- Velocity weirdness
- Surface gravity (log g)
- Effective temperature (Teff)
- Metallicity (M/H)
- Alpha-element enhancements (alpha/M)
- Carbon abundance (C/M)
- Nitrogen abundance (N/M)

Plots are interactive:

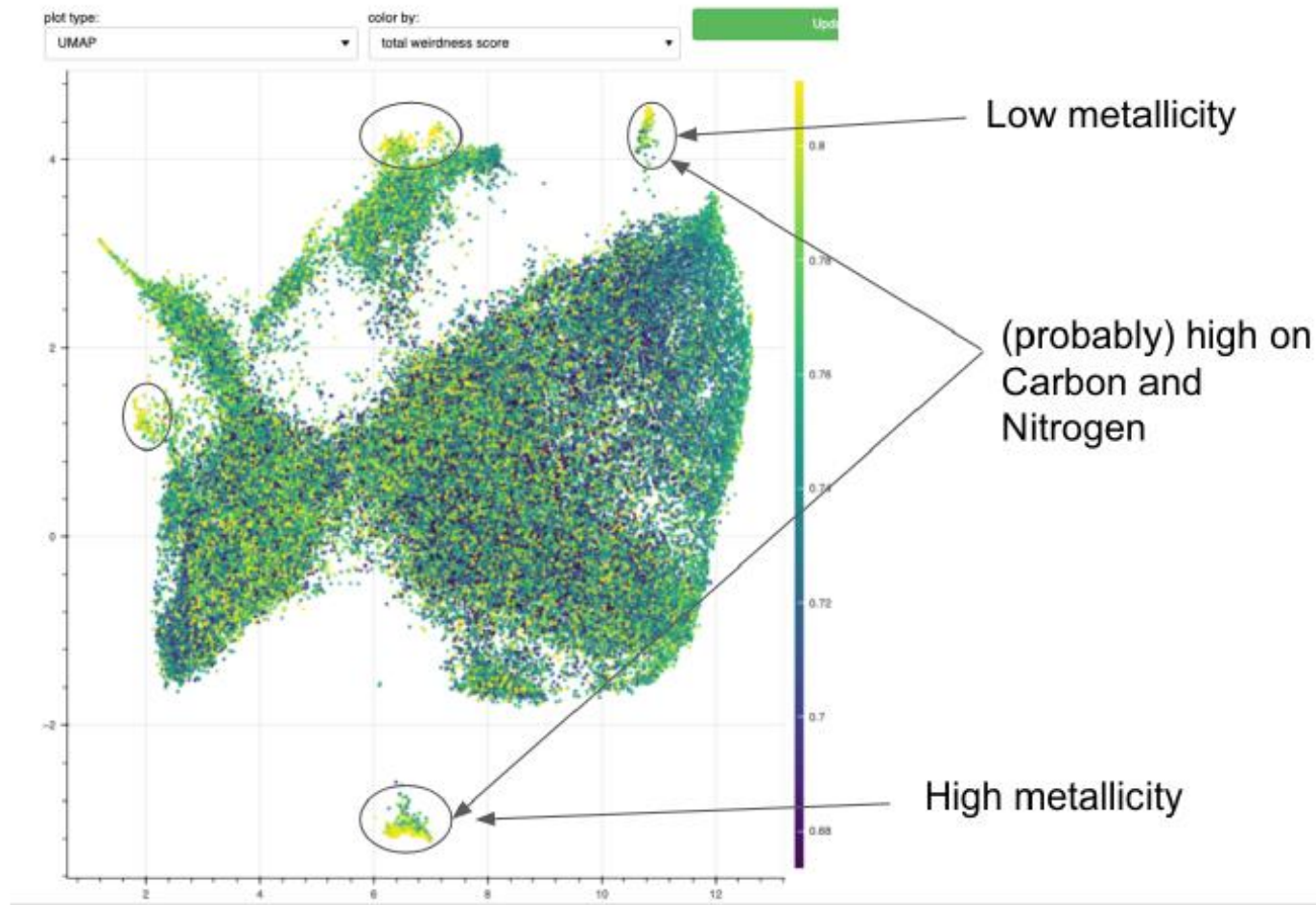
- Can select and inspect single objects
- See where they move between views
- Can select multiple and plot median spectrum



# EXPLORE SDA – S-Disco



UMAP + Z



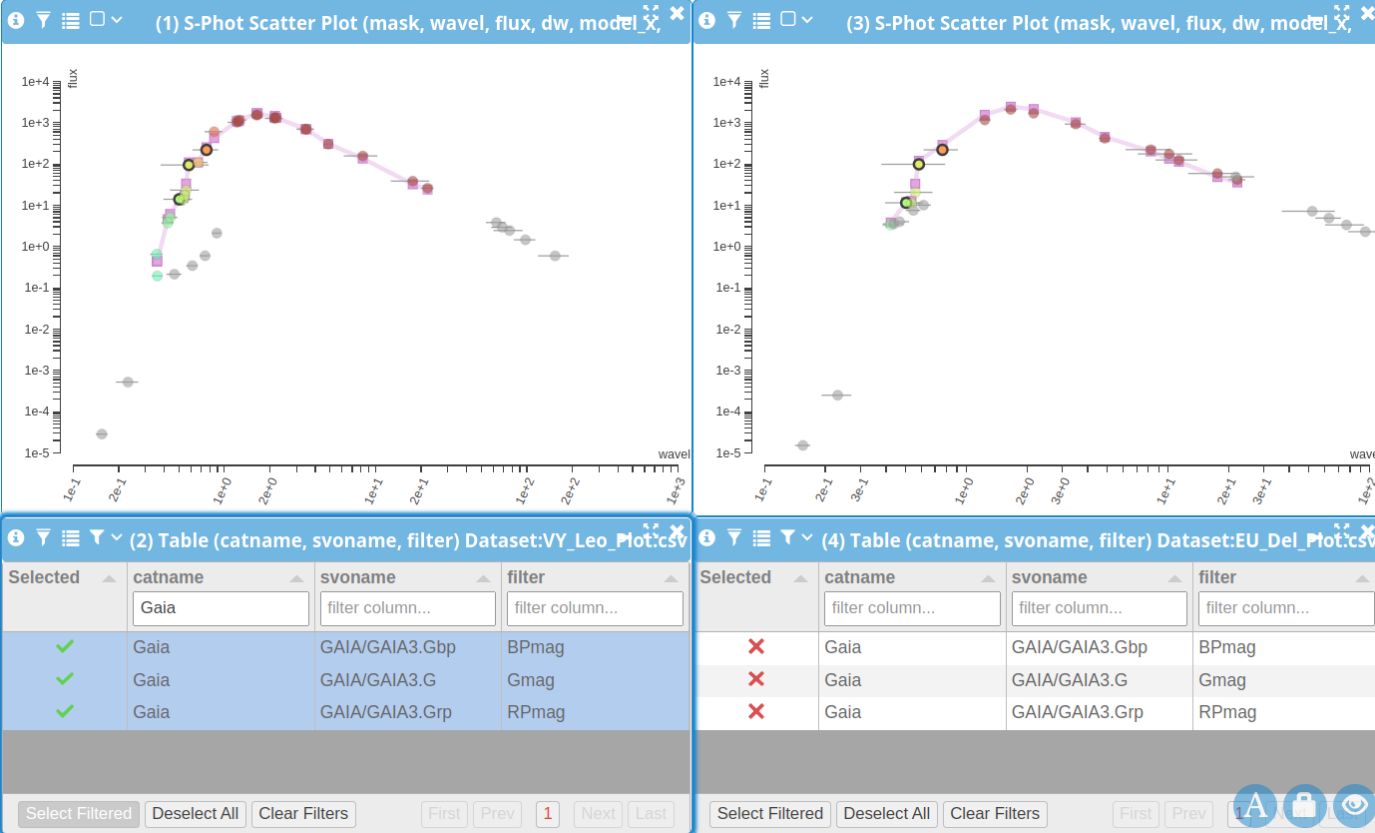
UMAP – weirdness score

# EXPLORE SDA – S-Phot



- New surveys are providing data on 10% of all stars in the Galaxy
- Allows population and evolution studies
  - Example Gaia: 1 billion accurate distance to stars
- Applications
  - Identifying stars with infrared or blue excess (star formation; activity; star death)
  - Mapping the solar neighbourhood
- Requirements
  - Determine accurate stellar parameters
  - Correlating many different databases with disparate information

# EXPLORE SDA – S-Phot



# EXPLORE SDA – S-Phot



The screenshot displays the EXPLORE SDA interface with four S-Phot scatter plots. A 'Create Annotation' dialog box is open, allowing users to create annotations for visualizations. The dialog includes fields for visualization selection, title, and description, as well as read and write access controls.

**Table 1: Categorical Data**

Selected	catname	svoname
	Gaia	

**Table 2: Numerical Data**

Selected	catname	svoname	filter
✓	Gaia	GAIA	
✓	Gaia	GAIA	
✓	Gaia	GAIA	

**Table 3: Filtered Data**

Selected	catname	svoname	filter
✗	Akari/FIS	AKARI/FIS.WIDE-S	S90
✗	Herschel/70	Herschel/PACS.blue	Flux
✗	Herschel/100	Herschel/PACS.green	Flux
✗	Herschel/160	Herschel/PACS.red	Flux

**Create Annotation Dialog**

Select Visualization to Annotate: (1) S-Phot Scatter Plot (mask, wavel, flux, dw, model\_x...)

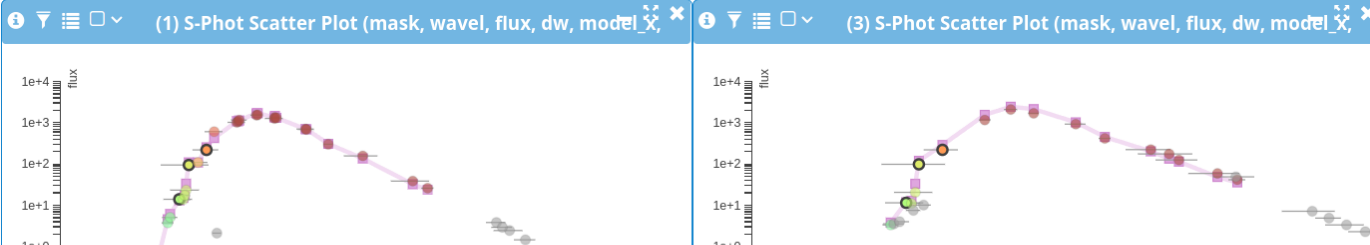
Title: Outliers in VY\_Leo

Description: Outliers detected with PySSED algo need further investigation.

Read Access: Public | Write Access: Public



# EXPLORE SDA – S-Phot



**Create Annotation**

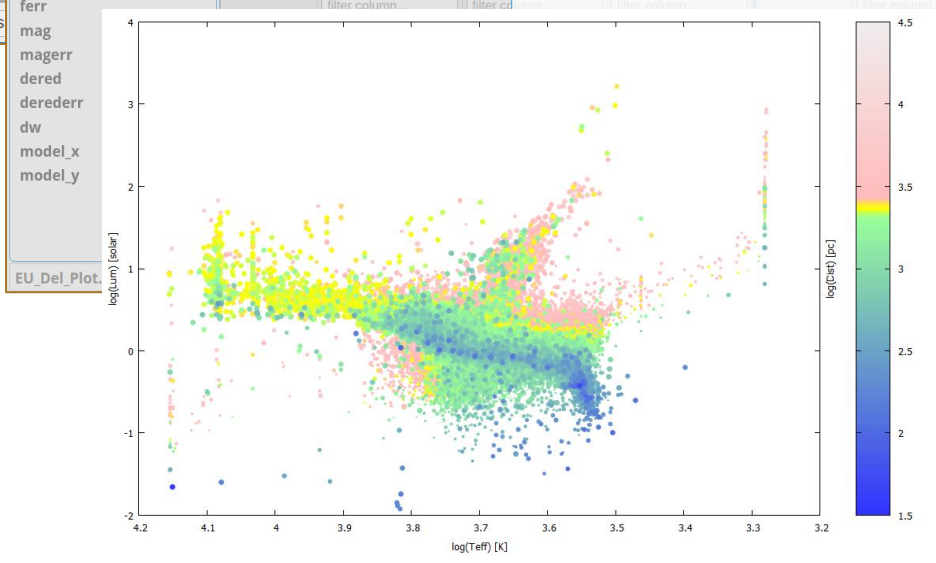
Select Visualization to Annotate  
(1) S-Phot Scatter Plot (mask, wavel, flux, dw, model\_x...)

Title  
Outliers in VY\_Leo

Description  
Outliers detected with PySSED algo need further investigation.

Select Filtered Deselect All Clear Filters

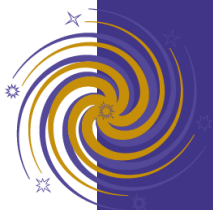
Selected	catname	svoname
	Gaia	
✓	Gaia	GAIA
✓	Gaia	GAIA
✓	Gaia	GAIA



**synchronisation**

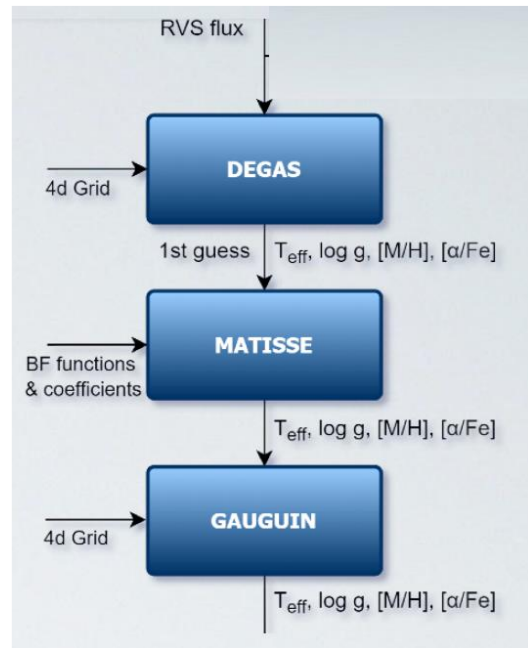
country: India  
life-expectancy: 66.54545454545455

- Argentina
- Australia
- Austria
- Brazil
- Canada
- China
- European Union
- France
- Germany
- India
- Indonesia
- Italy
- Japan
- Mexico
- Russia
- Saudi Arabia
- South Africa
- South Korea
- Turkey
- United Kingdom
- United States

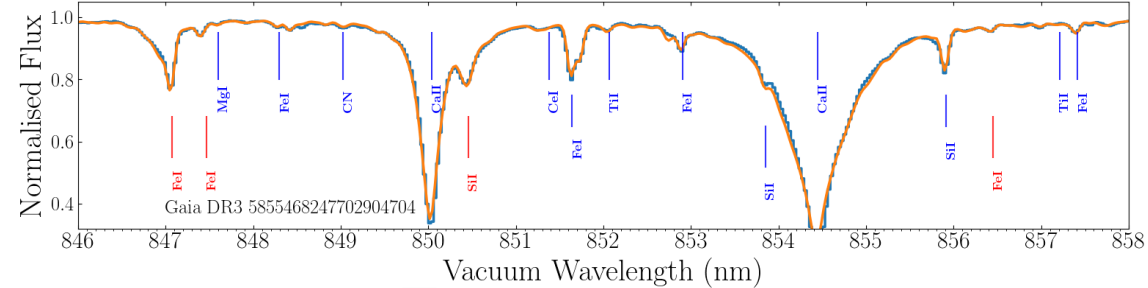


# EXPLORE SDA – G-Arch

Stellar parameter fit and abundance retrieval



From Recio-Blanco et al. (2022)



EXPLORE G-Arch

Configuration Panel

Processing completed

Save/export results

Results saved to workspace

Visualisation parameters

X:  $T_{\text{eff}}$  Y:  $\log g$  Z:  $[\text{Fe}/\text{H}]$

Visualisation selection

Sun

Arcturus

Analysis

Search siblings Galactic model

Detect outliers

Fit details

Dust extinction

Work in progress:

- Spectral similarity finder
- Inference of parameters for “Failed Fits”

# Invitation to join as beta user and get early access

- Looking for few extra beta users for each SDA
- In-person at EAS on June 29<sup>th</sup> (afternoon)
- Contact us at [contact@explore-platform.eu](mailto:contact@explore-platform.eu)



# EXPLORE

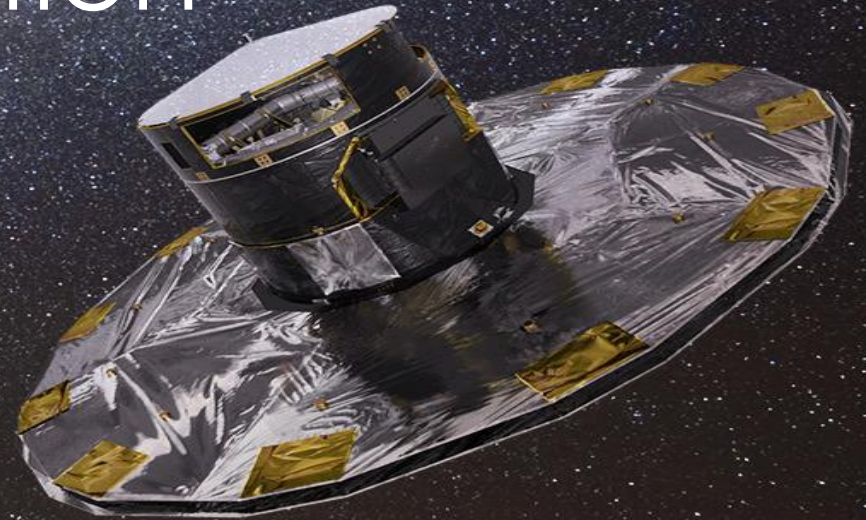
Thank you for your attention

More info on:

[explore-platform.eu](https://explore-platform.eu)

Get in touch:

[contact@explore-platform.eu](mailto:contact@explore-platform.eu)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under agreement No 101004214

