



eur^oPLANET2024
Research Infrastructure

GMAP
Geological Mapping

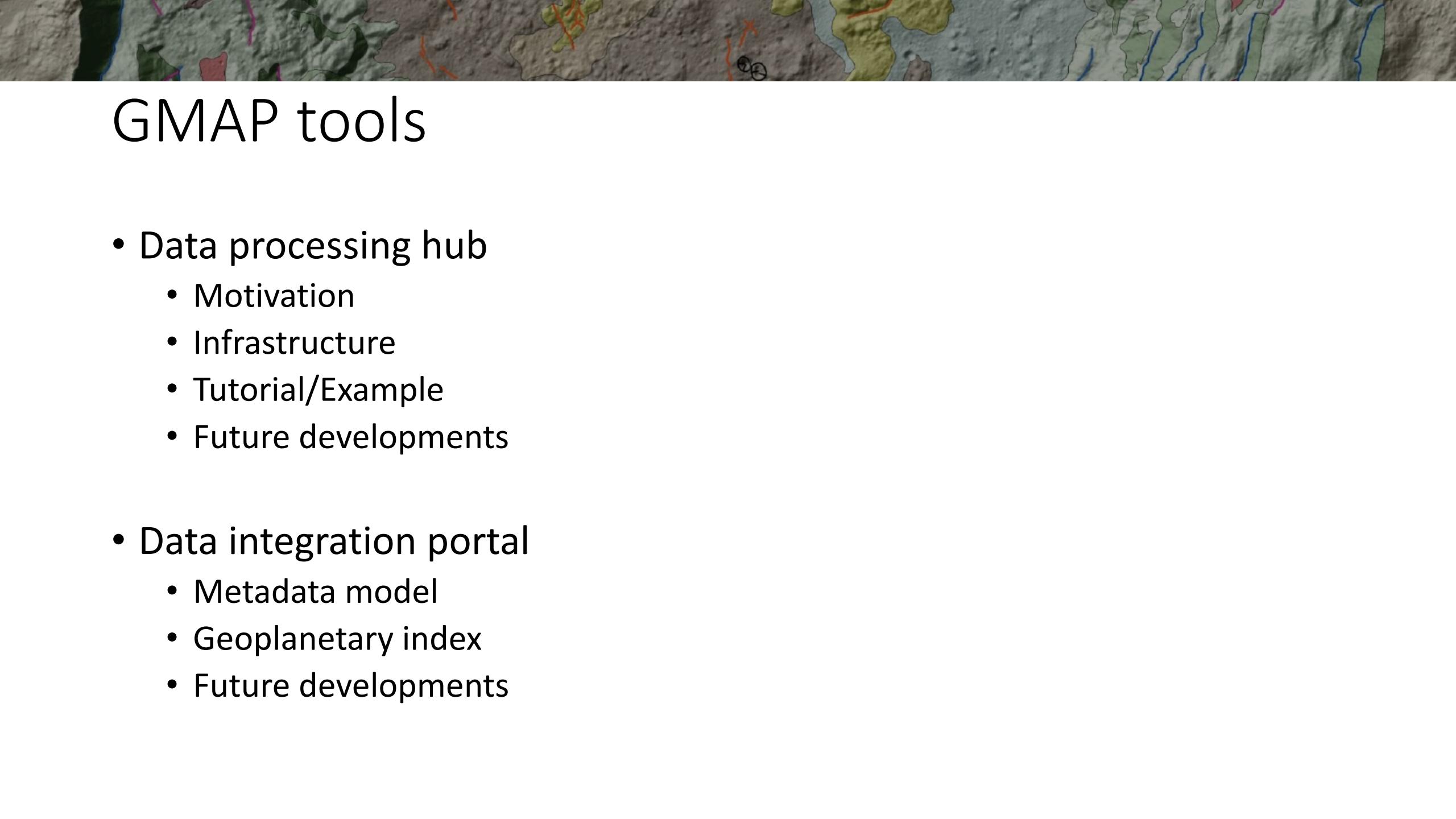
GMAP tools

Carlos H. Brandt, Angelo P. Rossi, Giacomo Nodjoumi, and the GMAP team
cbrandt@constructor.university

C>ONSTRUCTOR
UNIVERSITY



Europlanet 2024 RI has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871149



GMAP tools

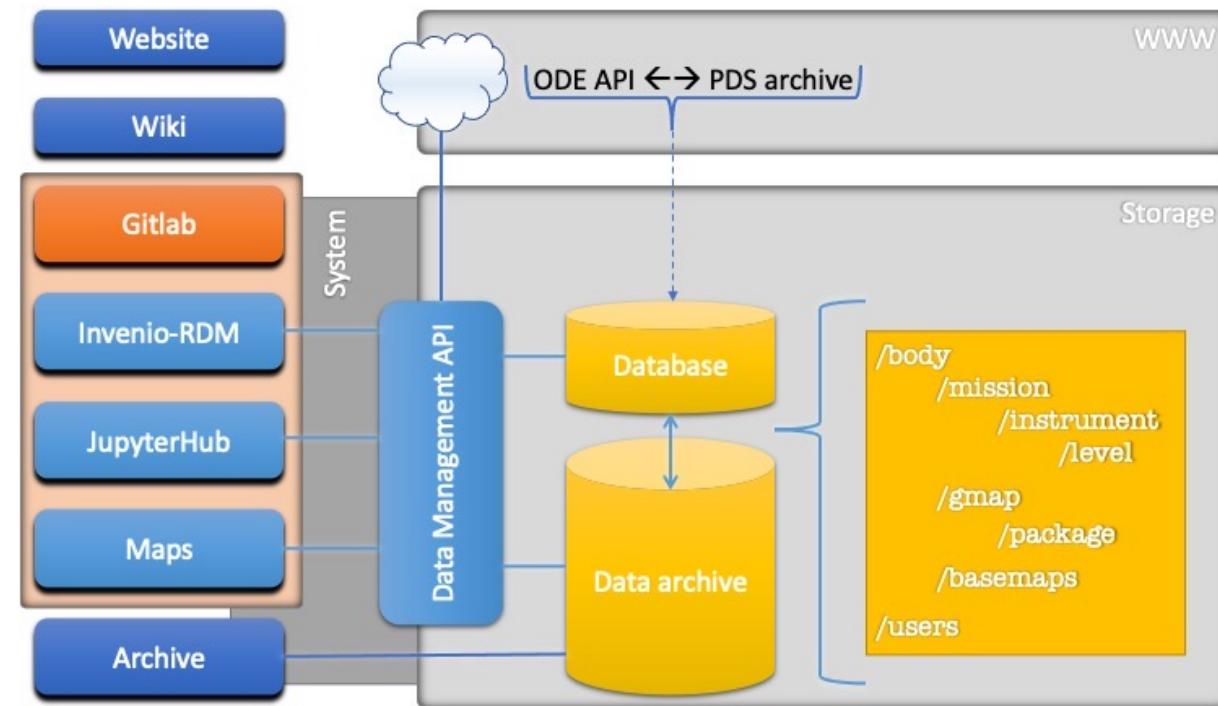
- Data processing hub
 - Motivation
 - Infrastructure
 - Tutorial/Example
 - Future developments
- Data integration portal
 - Metadata model
 - Geoplanetary index
 - Future developments



Presentation Goals

- Offer you a ready-to-use planetary data analysis environment
- Motivate your contribution – use, features request, coding
- Show off our awesome work ;)

Infrastructure



Infrastructure

GMAP services/interfaces

Website

Wiki

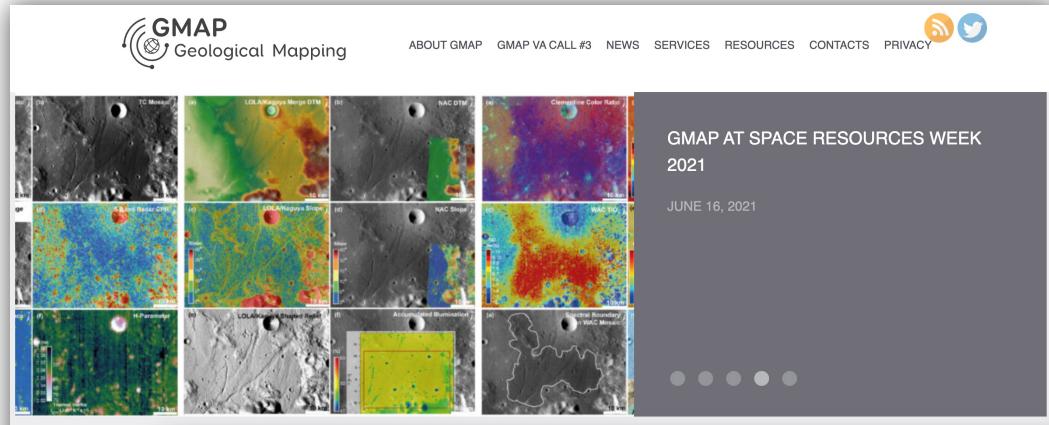
Gitlab

Invenio-RDM

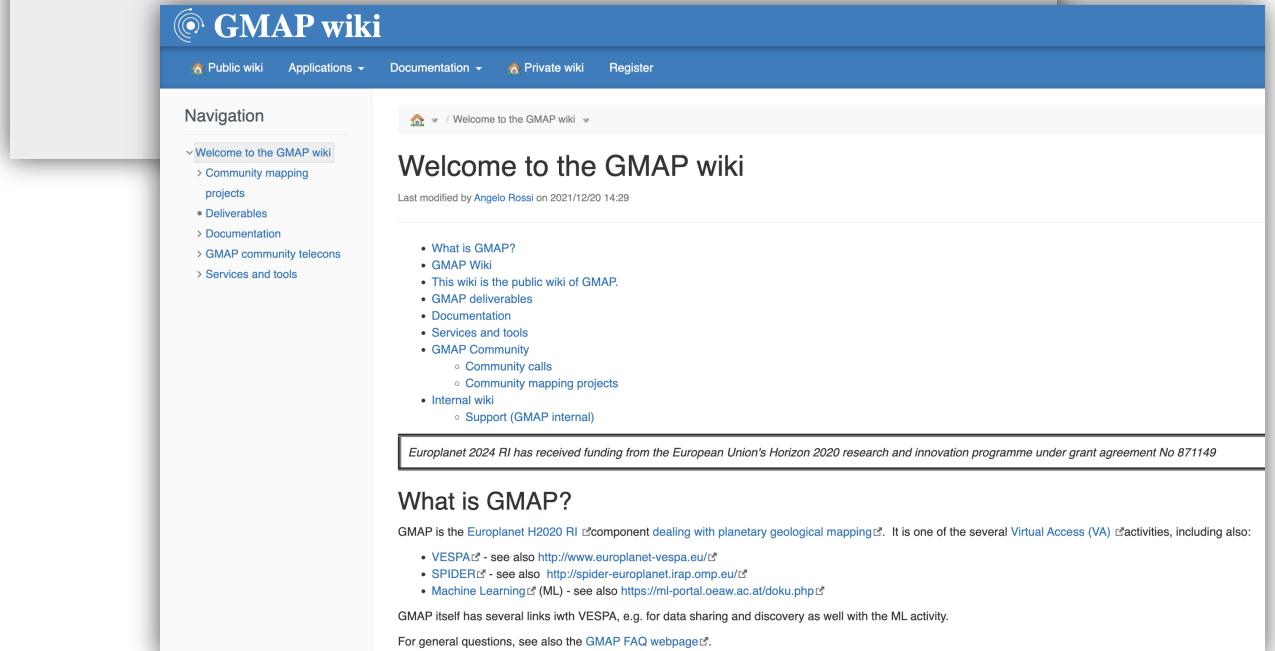
JupyterHub

Infrastructure – Info & Documentation

- Website
 - <https://europlanet-gmap.eu>



The screenshot shows the GMAP website homepage. At the top, there's a navigation bar with links to 'ABOUT GMAP', 'GMAP VA CALL #3', 'NEWS', 'SERVICES', 'RESOURCES', 'CONTACTS', and 'PRIVACY'. Below the navigation is a grid of nine geological maps labeled (a) through (i). To the right of the maps is a dark grey box containing the text 'GMAP AT SPACE RESOURCES WEEK 2021' and 'JUNE 16, 2021'. A series of five small circular dots is positioned below the dark box.



The screenshot shows the GMAP wiki homepage. The header includes a logo, a search bar, and navigation links for 'Public wiki', 'Applications', 'Documentation', 'Private wiki', and 'Register'. The main content area has a blue sidebar titled 'Navigation' with sections for 'Welcome to the GMAP wiki', 'Community mapping projects', 'Deliverables', 'Documentation', 'GMAP community telecons', and 'Services and tools'. The main content area features a 'Welcome to the GMAP wiki' section with a 'Last modified by Angelo Rossi on 2021/12/20 14:29' message. It also lists 'What is GMAP?' and other project details. A funding acknowledgment at the bottom states: 'Europlanet 2024 RI has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871149'. The footer contains a 'What is GMAP?' section, a note about the project's role in the Europlanet H2020 RI, and links to related projects like VESPA, SPIDER, and Machine Learning.

Infrastructure – Project Management

- GitLab
 - <https://git.europlanet-gmap.eu>
 - Groups
 - Repositories
 - Code prototypes
 - Working documents
 - Issues
 - Discussions
 - Actions
 - User authentication
 - Jupyter-Hub
 - Data Portal(!)

The screenshot shows the GitLab interface for the 'GMAP' group. The left sidebar displays group statistics: Group ID: 12, 27 issues, 1 merge request, and 1 Kubernetes package. The main content area shows a list of subgroups and projects. Subgroups include 'data-access', 'data-packages', 'data-publication', 'community support', 'csfd', 'gmap', 'GMAP community documentation', 'gmap_coordination', 'Gmap Website Docker', 'gmap-website-wordpress-docker', and 'GMAP Winter School 2024'. Projects listed under 'Subgroups and projects' include 'data-access', 'data-packages', 'data-publication', 'community support', 'csfd', 'gmap', 'GMAP community documentation', 'gmap_coordination', 'Gmap Website Docker', 'gmap-website-wordpress-docker', and 'GMAP Winter School 2024'. The 'data-access' project is described as having overlapping components with 'data-publication'. The 'GMAP community documentation' project is described as an authoring location for other frontends like web pages and sphinx. The 'gmap' project is a placeholder for common tasks. The 'GMAP Winter School 2024' project is described as organizing the winter school materials.

Project	Description
data-access	Data access interfaces, graphical or not, software components. This group overlaps with data-publication
data-packages	Placeholder for the projects implementing data pipelines. Gitlab workflows, web interface for data submission, package validator, etc.
data-publication	Group for the projects implementing data pipelines. Gitlab workflows, web interface for data submission, package validator, etc.
community support	Placeholder for the projects implementing data pipelines. Gitlab workflows, web interface for data submission, package validator, etc.
csfd	Crater Size Frequency distribution basic plotting tool (jupyter, qgis plugin, tbd), using what is available, e.g. Riedel CSDD, targeting m...
gmap	Placeholder to issue that have no group/project yet. Or code snippets for common tasks not restricted to one or another working group.
GMAP community documentation	Authoring location and source for other frontends (e.g. web pages, sphinx, etc.) of GMAP-relevant community documentation, for pip...
gmap_coordination	issue and action tracking for GMAP JRA, VA
Gmap Website Docker	gmap website
gmap-website-wordpress-docker	the gmap website dockerized
GMAP Winter School 2024	Repo for organising the Winter School, including materials. When mature, materials will go on the relevant GitHub repository. Data will go, as customary, on the cloud and in due time on Zenodo for longer-term availability

Infrastructure – processing & data publication

- **Jupyter-Hub**

- Interactive data analysis
- Pre-set standard software
- Collaborative environment

- **Data Integration Portal**

- Standard data publication
 - Standard metadata
 - Standard interface
- Planetary datasets index

The collage illustrates the GMAP (Geological Mapping) infrastructure, which includes a Jupyter Hub for reproducible processing and a Data Integration Portal for standard data publication.

Jupyter-Hub Screenshot:

- File menu: File, Edit, View, Run, Kernel, Tabs, Settings, Help.
- Launcher tab: pds_data_load.ipynb
- File tree: / shared /
- Content pane: Europlanet GMAP @ ERIM 2023, PDS data load/view, text about the notebook's purpose, and a note about using markdown.

Data Integration Portal Screenshot:

- Header: GMAP Geological Mapping, Communities, My dashboard, Log in, Sign up.
- Search bar: Search
- Content area:
 - Search Data: Use free-text search and tags to filter records, search page button.
 - Data Analysis: Deeper data inspection and processing tools, user environment button.
 - Documentation: Material about our data and map-making, wiki button.

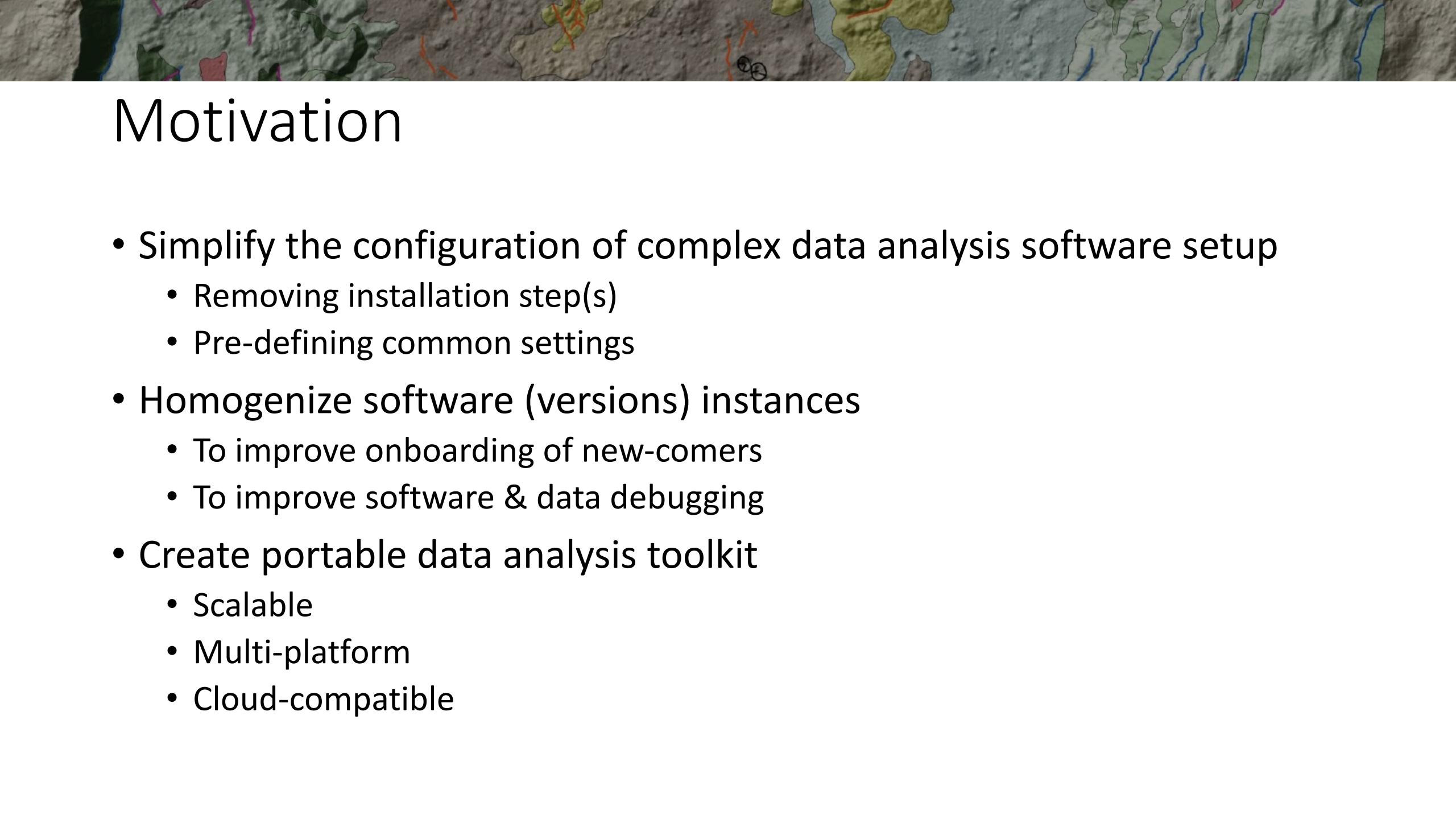


eur^oPLANET2024
Research Infrastructure

GMAP
Geological Mapping

GMAP Jupyter-Hub

Data Processing Hub

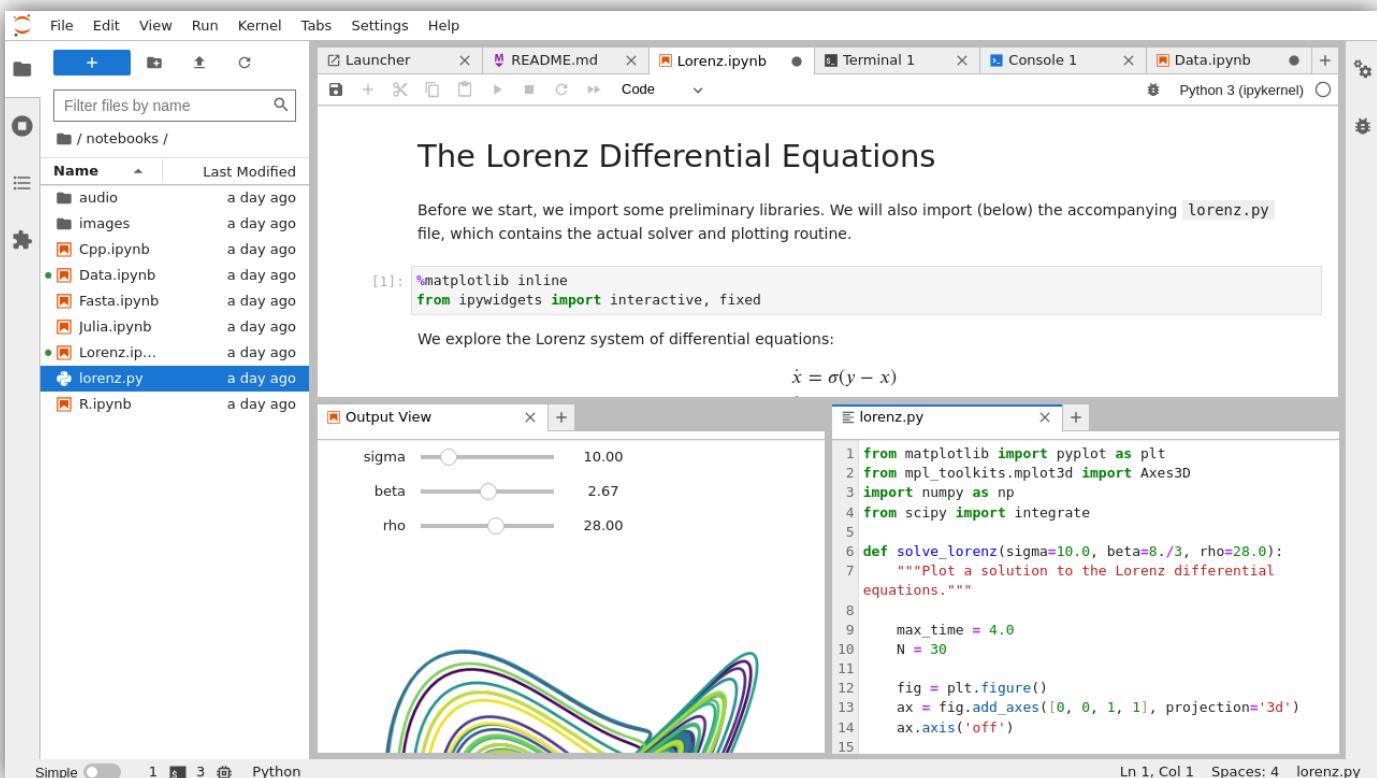


Motivation

- Simplify the configuration of complex data analysis software setup
 - Removing installation step(s)
 - Pre-defining common settings
- Homogenize software (versions) instances
 - To improve onboarding of new-comers
 - To improve software & data debugging
- Create portable data analysis toolkit
 - Scalable
 - Multi-platform
 - Cloud-compatible

What is Jupyter-Hub?

- Jupyter-Hub is a manager of Jupyter (Lab/Notebook) instances
 - “Lab” == the modern version of “Notebook”
(say “Notebook”, use “Lab”)
- Handles user authentication
- Start/Stop user instances



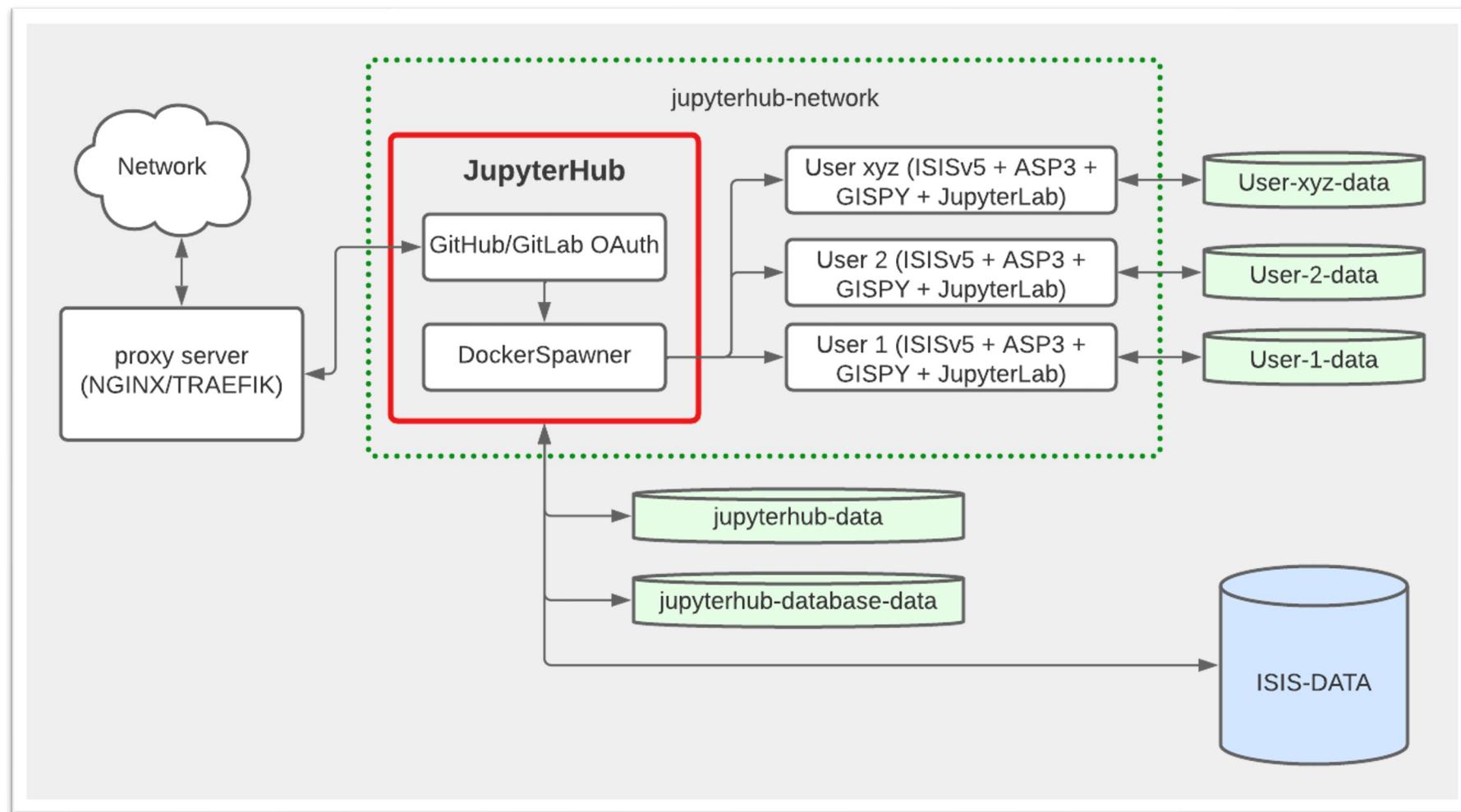
The screenshot shows the Jupyter-Hub interface with several tabs at the top: Launcher, README.md, Lorenz.ipynb, Terminal 1, Console 1, Data.ipynb, and Python 3 (ipykernel). The main area displays a notebook titled "The Lorenz Differential Equations". The code cell contains:

```
%matplotlib inline
from ipywidgets import interactive, fixed
```

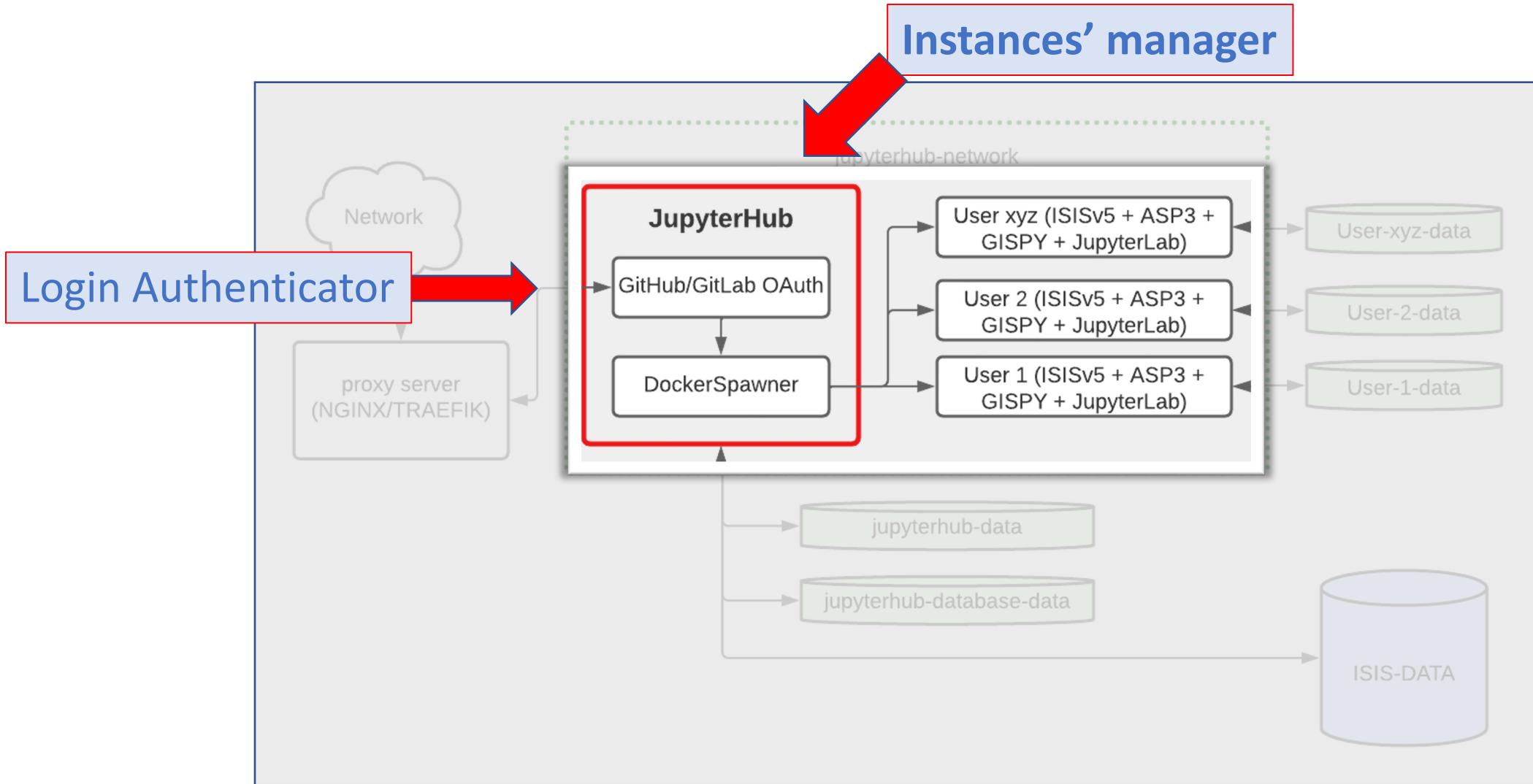
The text below the code says: "We explore the Lorenz system of differential equations:" followed by the mathematical equation $\dot{x} = \sigma(y - x)$. On the left, there is an "Output View" panel with sliders for sigma (10.00), beta (2.67), and rho (28.00). On the right, the "lorenz.py" file is shown with its code:1 from matplotlib import pyplot as plt
2 from mpl_toolkits.mplot3d import Axes3D
3 import numpy as np
4 from scipy import integrate
5
6 def solve_lorenz(sigma=10.0, beta=8./3, rho=28.0):
7 """Plot a solution to the Lorenz differential
 equations."""
8
9 max_time = 4.0
10 N = 30
11
12 fig = plt.figure()
13 ax = fig.add_axes([0, 0, 1, 1], projection='3d')
14 ax.axis('off')

At the bottom, it says "Ln 1, Col 1 Spaces: 4 lorenz.py".

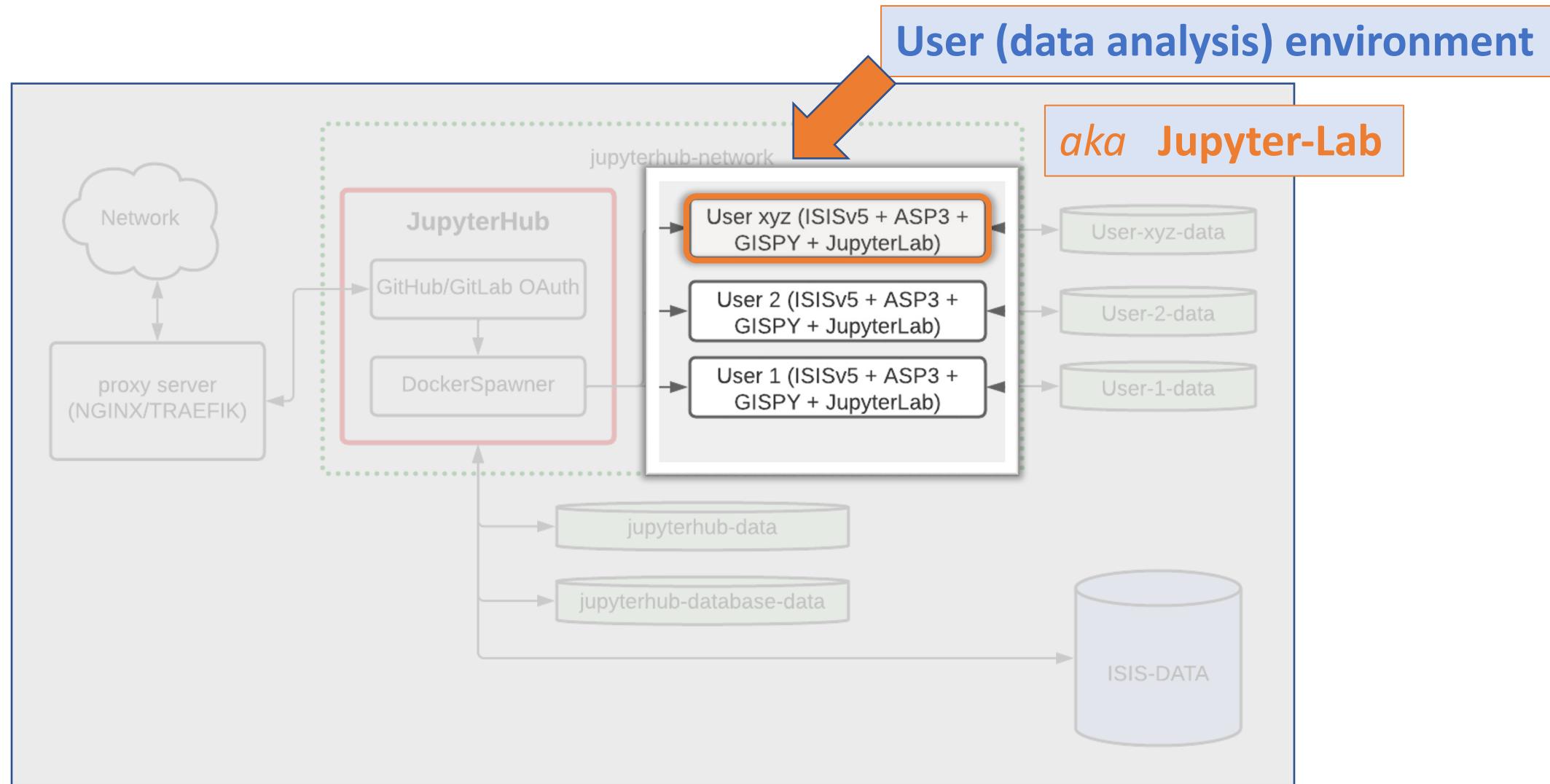
GMAP Jupyter-Hub



GMAP Jupyter-Hub

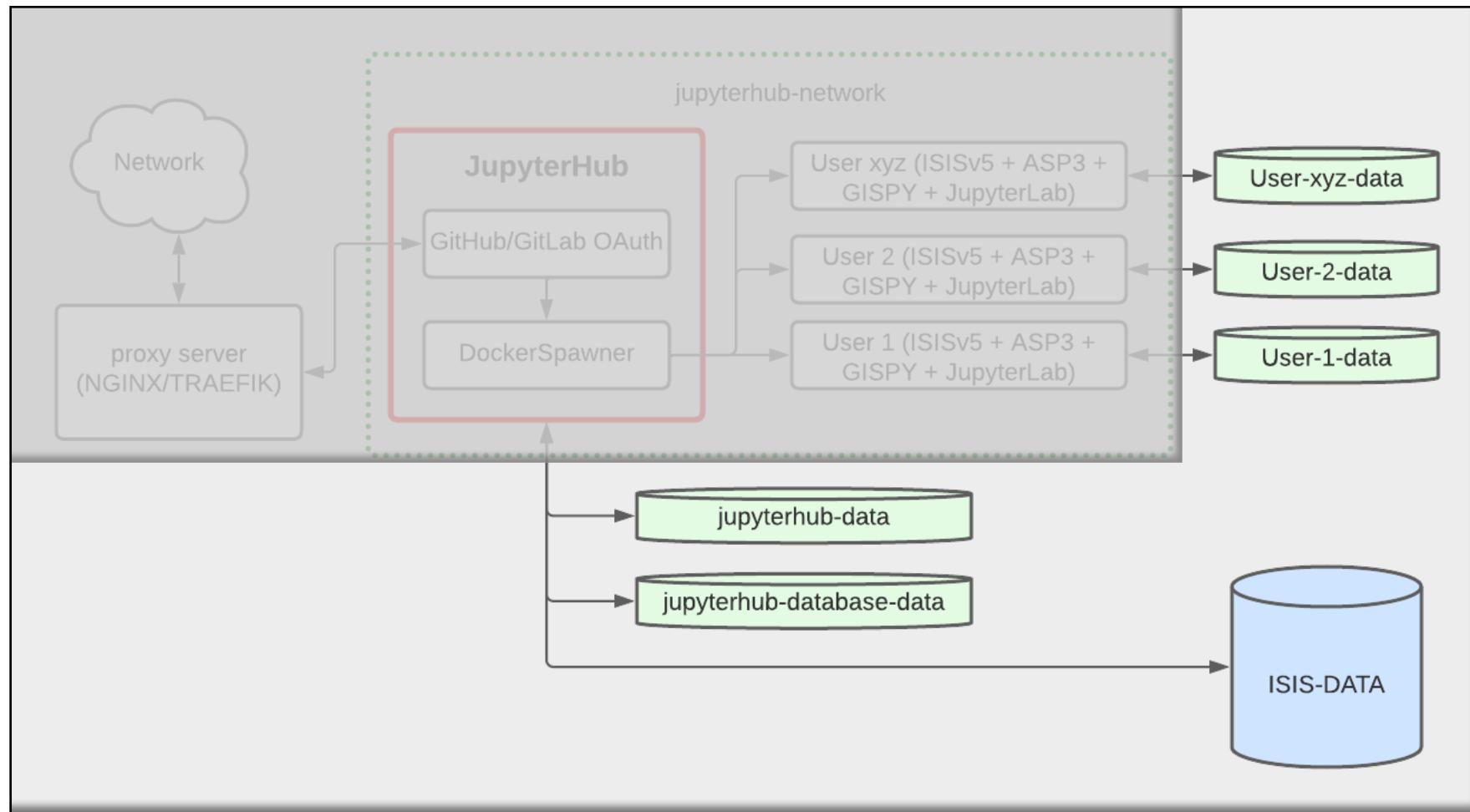


GMAP Jupyter-Hub



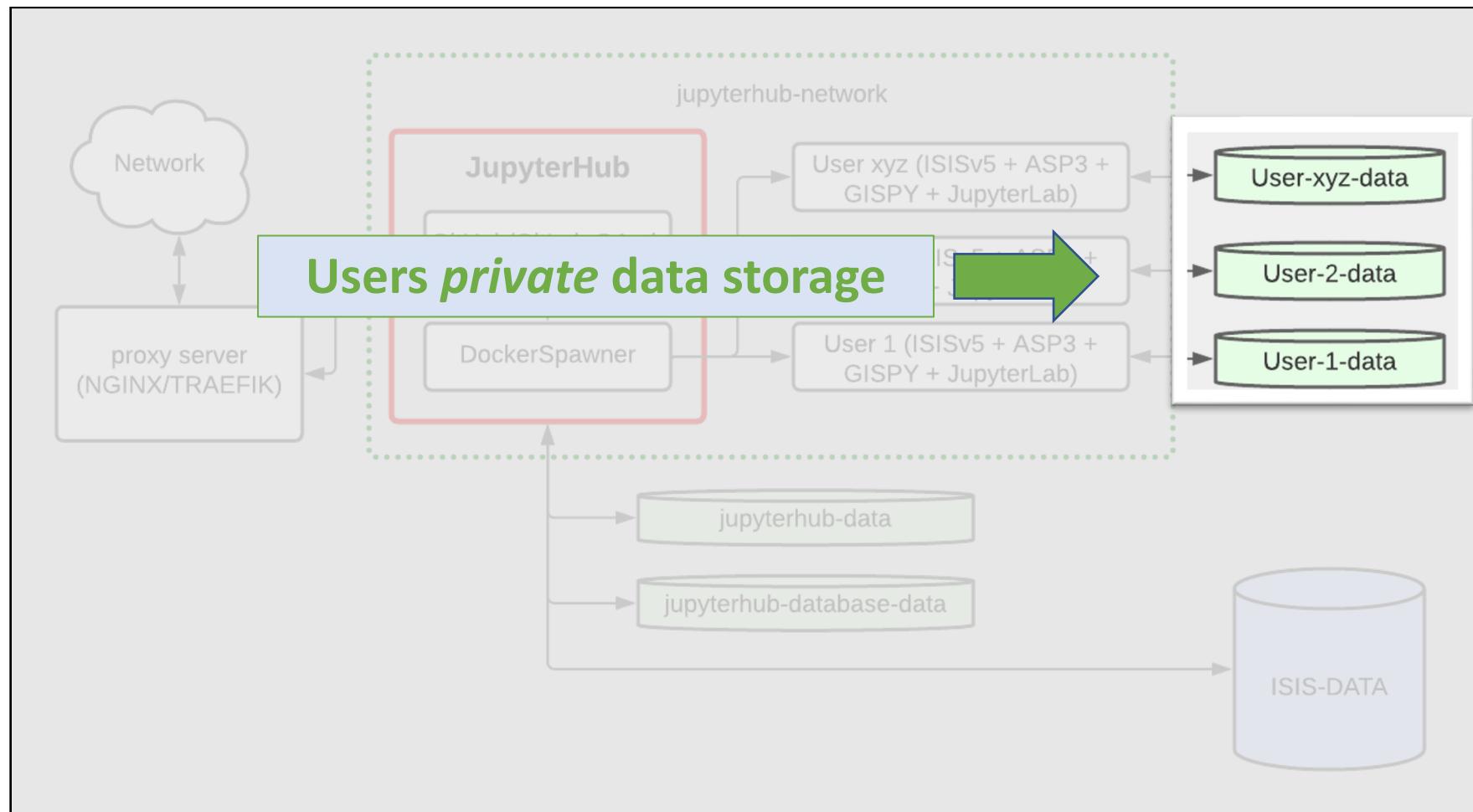
GMAP Jupyter-Hub

Persistent Data Volumes



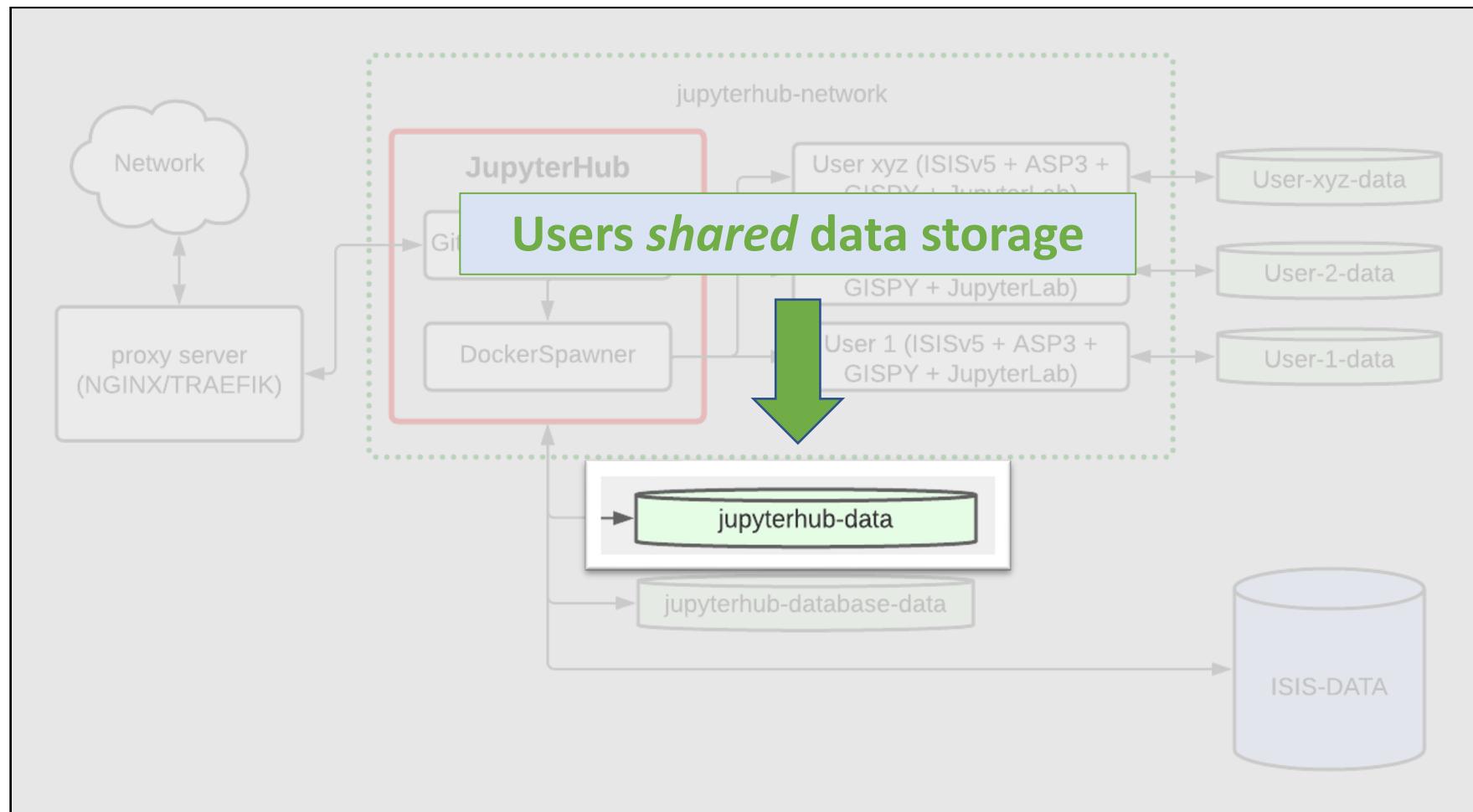
GMAP Jupyter-Hub

Persistent Data Volumes



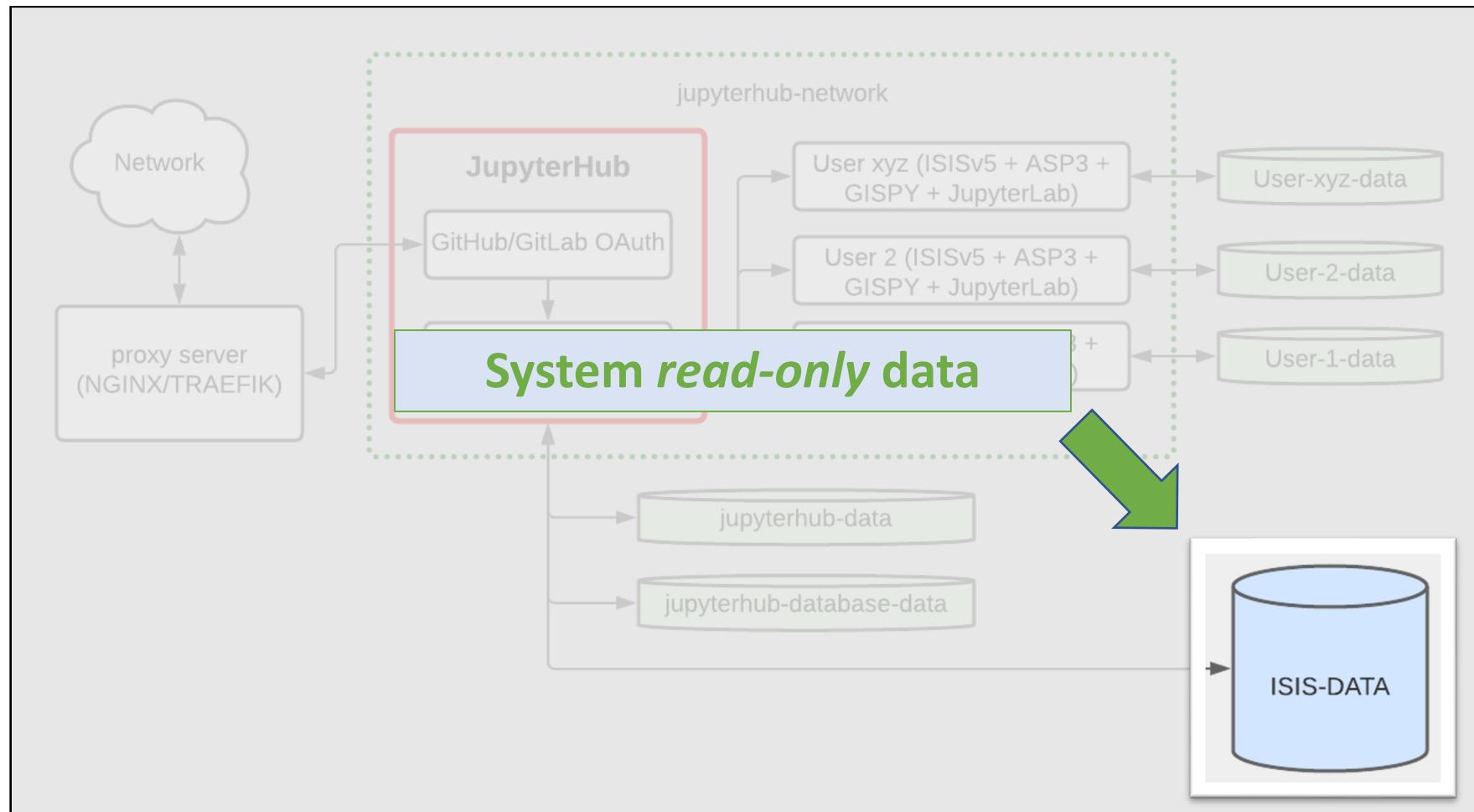
GMAP Jupyter-Hub

Persistent Data Volumes

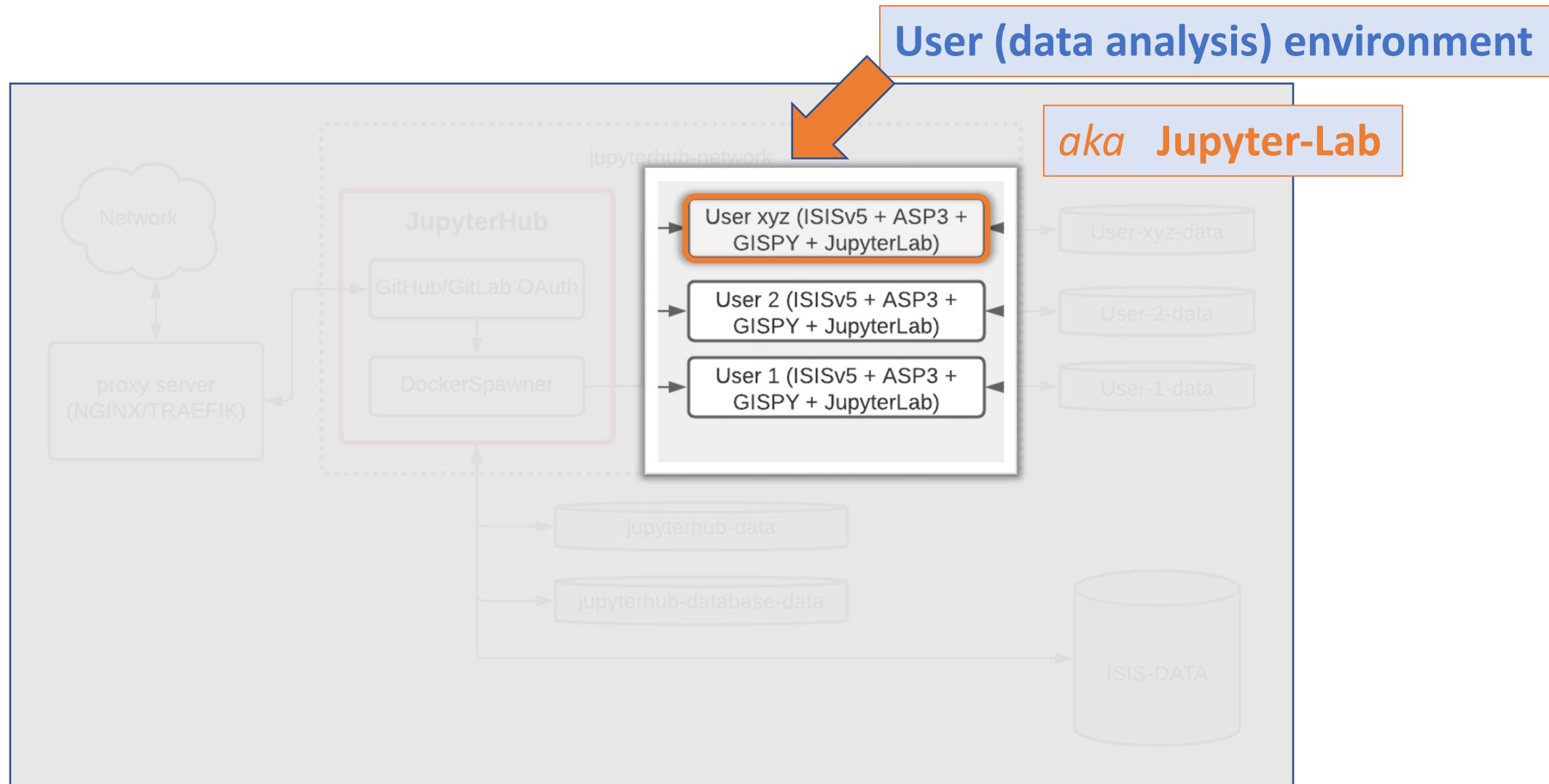


GMAP Jupyter-Hub

Persistent Data Volumes

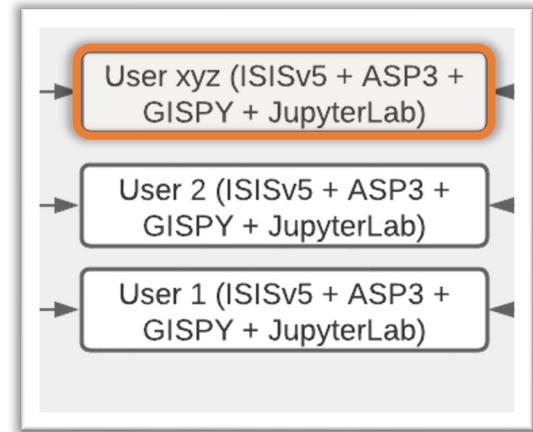


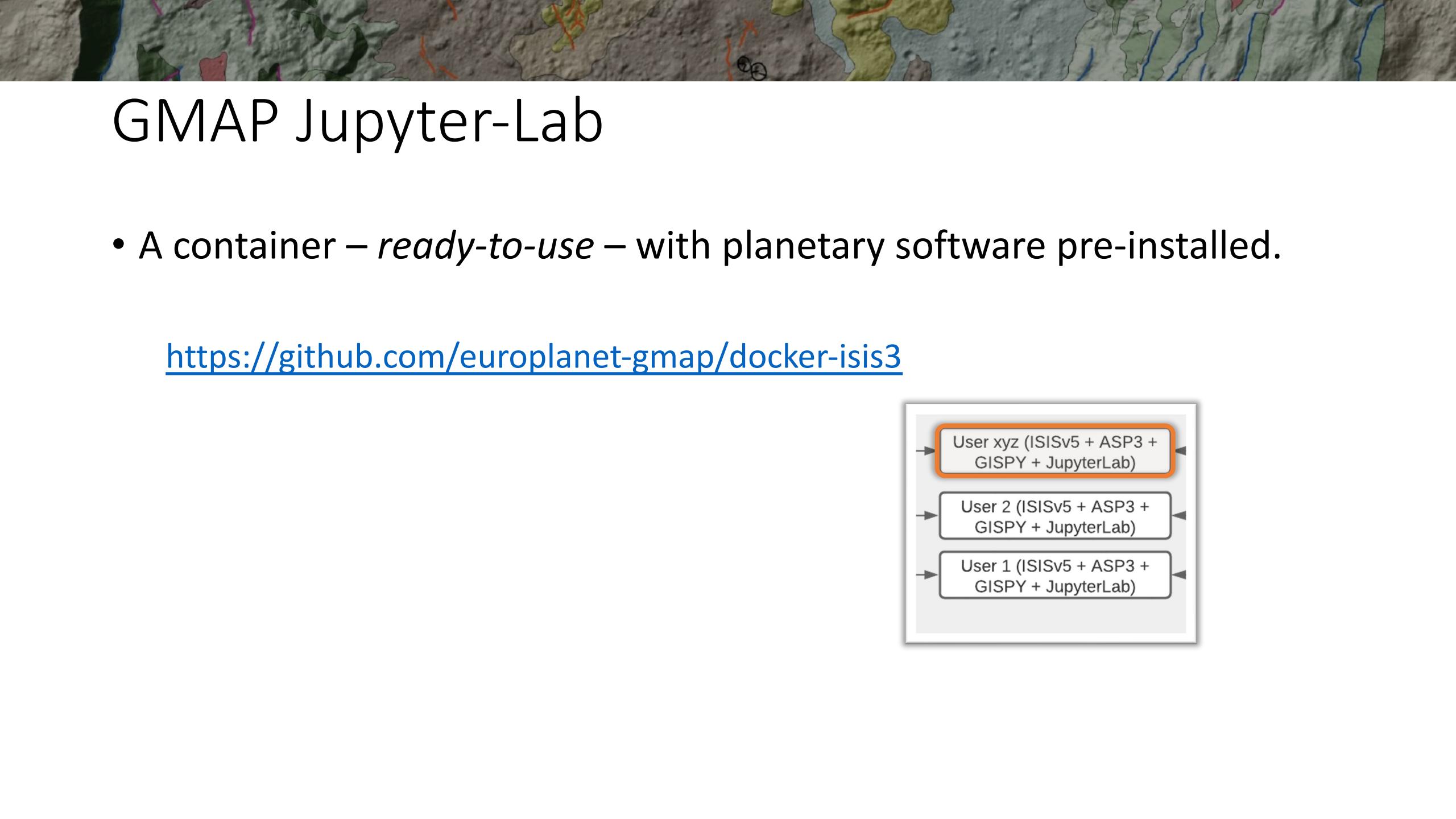
GMAP Jupyter-Lab



GMAP Jupyter-Lab

- A container – *ready-to-use* – with planetary software pre-installed.
 - ISIS (*Integrated Software for Imagers and Spectrometers*)
 - ASP (*Ames Stereo Pipeline*)
 - GDAL (*Geographical Data Abstraction Library*)
 - Python
 - geo/planetary data science tools
 - Geopandas
 - Shapely
 - Rasterio
 - etc.
 - Conda (package manager)

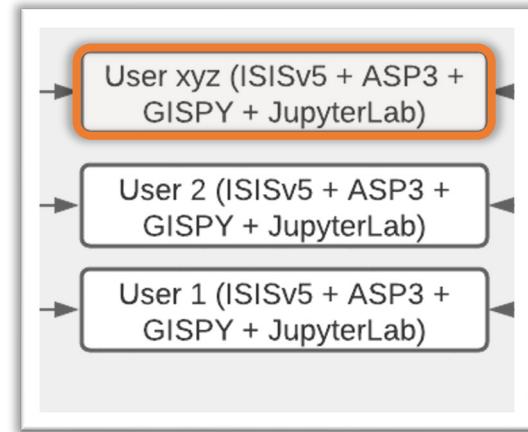


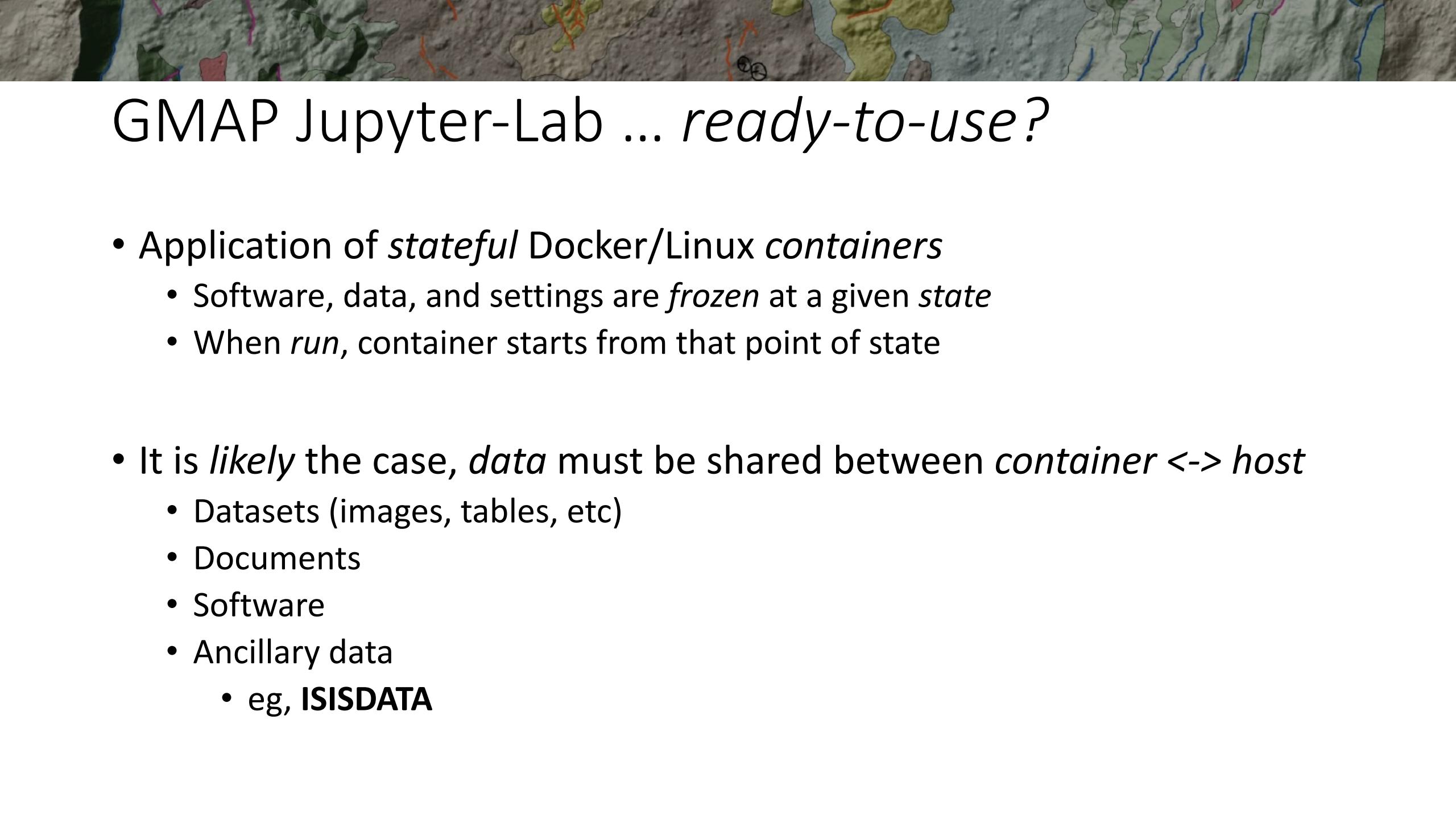


GMAP Jupyter-Lab

- A container – *ready-to-use* – with planetary software pre-installed.

<https://github.com/europlanet-gmap/docker-isis3>





GMAP Jupyter-Lab ... *ready-to-use?*

- Application of *stateful* Docker/Linux *containers*
 - Software, data, and settings are *frozen* at a given *state*
 - When *run*, container starts from that point of state
- It is *likely* the case, *data* must be shared between *container <-> host*
 - Datasets (images, tables, etc)
 - Documents
 - Software
 - Ancillary data
 - eg, **ISISDATA**



GMAP Jupyter-Lab ... *ready-to-use?*

OK, fine... but *how to use it* then...

- Install Docker (engine)
 - <https://docs.docker.com>
- *Pull* container image
 - gmap/jupyter-isis
 - <https://hub.docker.com/r/gmap/jupyter-isis>
 - gmap/jupyter-gispy
 - <https://hub.docker.com/r/gmap/jupyter-gispy>
- Download ISISDATA
 - <https://github.com/DOI-USGS/ISIS3/blob/dev/README.md#The-ISIS-Data-Area>

GMAP Jupyter-Lab ... *ready-to-use?*

OK, fine... but *how to use it* then...

1. Install Docker (engine)

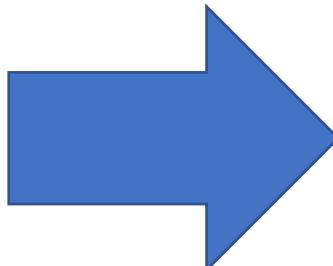
- <https://docs.docker.com>

2. Pull (container) image

- gmap/jupyter-isis
 - <https://hub.docker.com/r/gmap/jupyter-isis>
- gmap/jupyter-gispy
 - <https://hub.docker.com/r/gmap/jupyter-gispy>

3. Download ISISDATA

- <https://github.com/DOI-USGS/ISIS3/blob/dev/README.md#The-ISIS-Data-Area>



4. Run container (image)



GMAP Jupyter-Lab ... *ready-to-use?*

In more practical terms... *i.e.*, in clear commands:

- Consider Docker installed, data downloaded.

1. Pull image:

```
docker pull gmap/jupyter-gispy
```

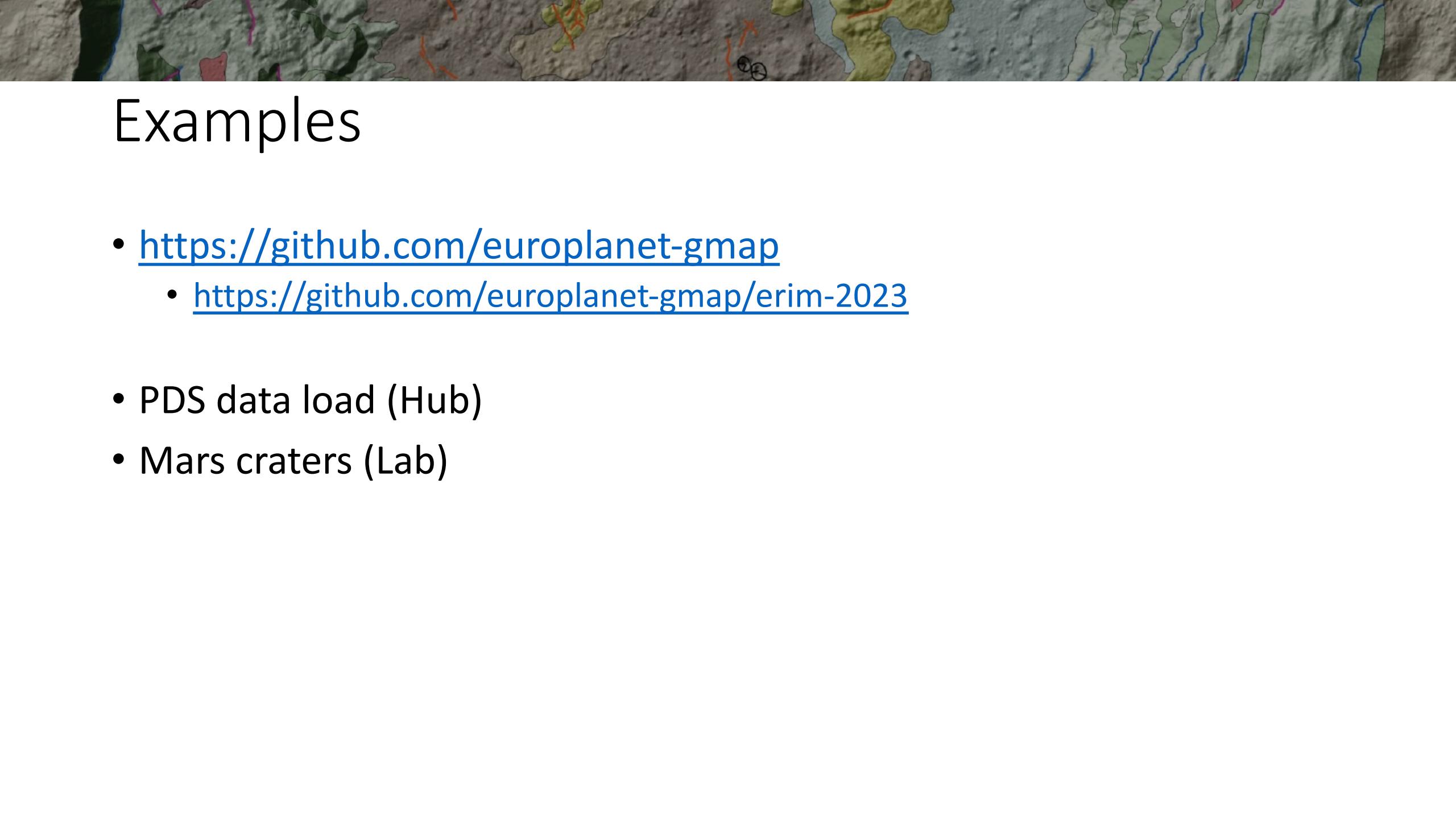


2. Run image:

```
docker run --rm -p 8888:8888 gmap/jupyter-gispy
```

3. Access Jupyter:

```
http://localhost:8888
```



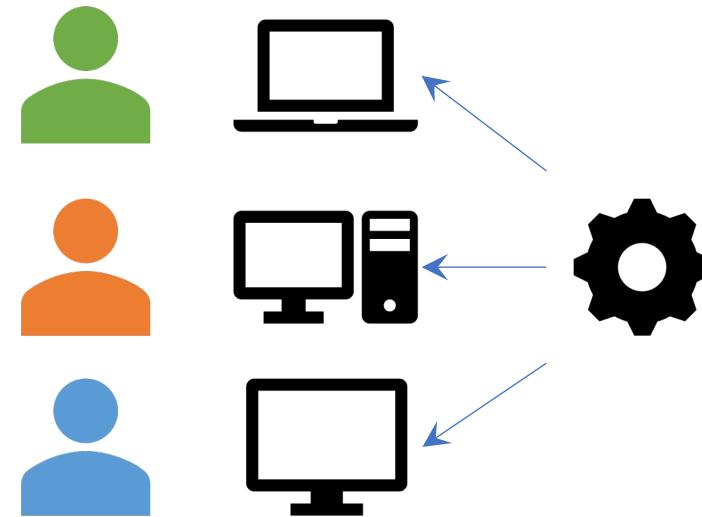
Examples

- <https://github.com/europlanet-gmap>
 - <https://github.com/europlanet-gmap/erim-2023>
- PDS data load (Hub)
- Mars craters (Lab)

A note on *portability*

Portability: software can run in different platforms:

- GNU/Linux, MacOS, Windows...



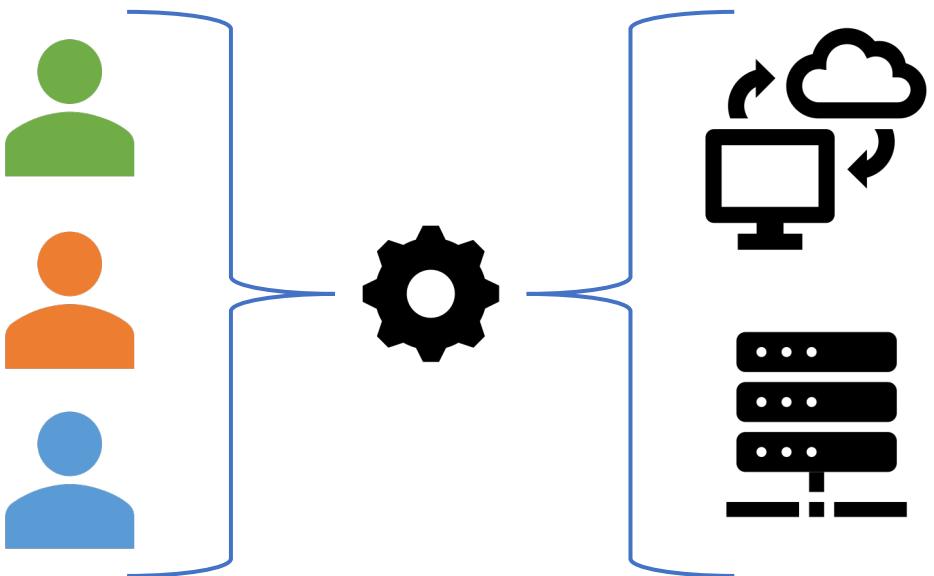
A note on *portability*

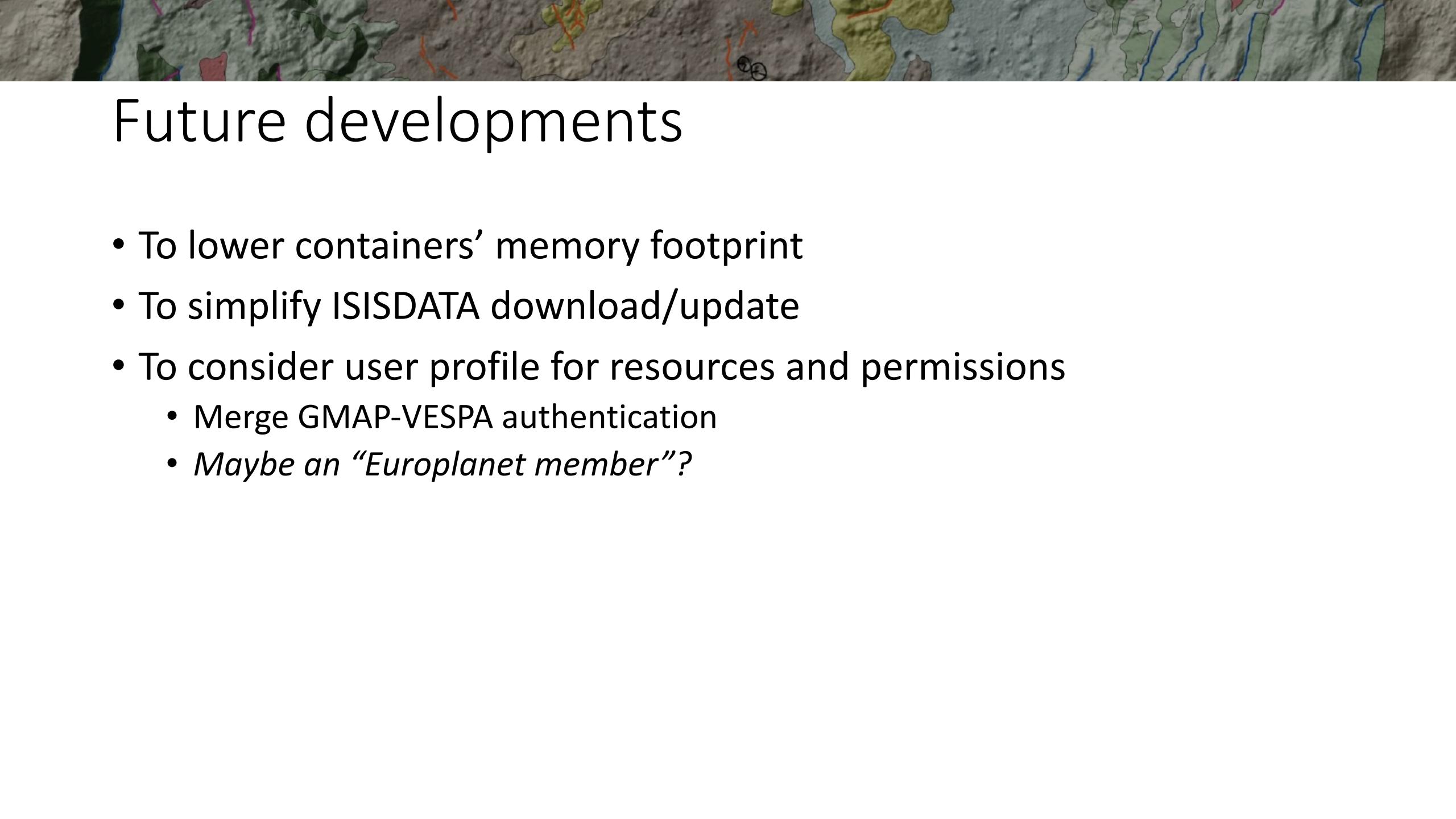
Portability: software can run in different platforms:

- GNU/Linux, MacOS, Windows...

but not only... *containers* allow us
to go beyond and *scale up*:

- Clusters and Cloud platforms
- ...the very same software, settings, data





Future developments

- To lower containers' memory footprint
- To simplify ISISDATA download/update
- To consider user profile for resources and permissions
 - Merge GMAP-VESPA authentication
 - *Maybe an “Europlanet member”?*

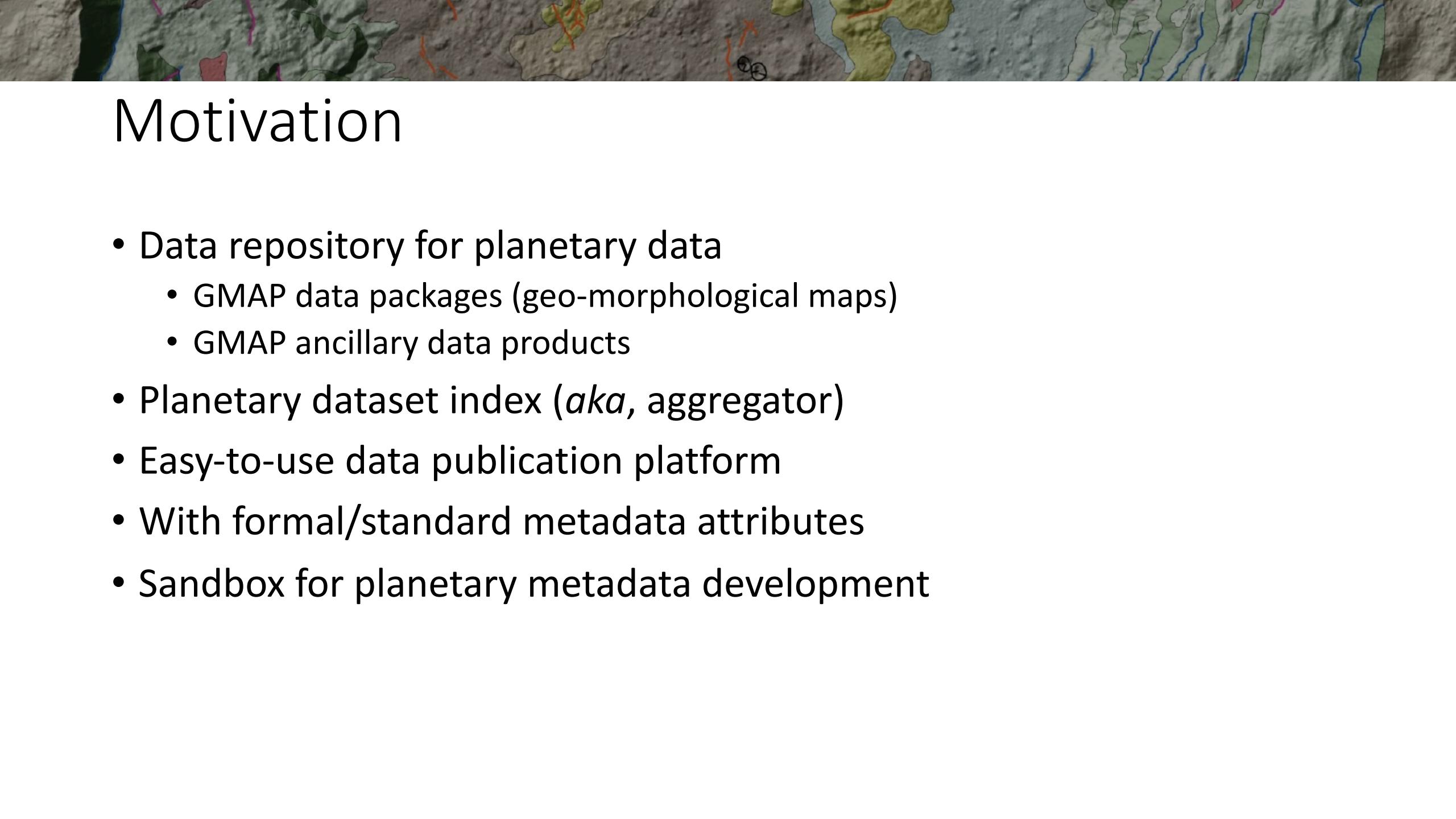


eur^oPLANET2024
Research Infrastructure

GMAP
Geological Mapping

Data Integration Portal

Data publication and planetary data index



Motivation

- Data repository for planetary data
 - GMAP data packages (geo-morphological maps)
 - GMAP ancillary data products
- Planetary dataset index (*aka*, aggregator)
- Easy-to-use data publication platform
- With formal/standard metadata attributes
- Sandbox for planetary metadata development

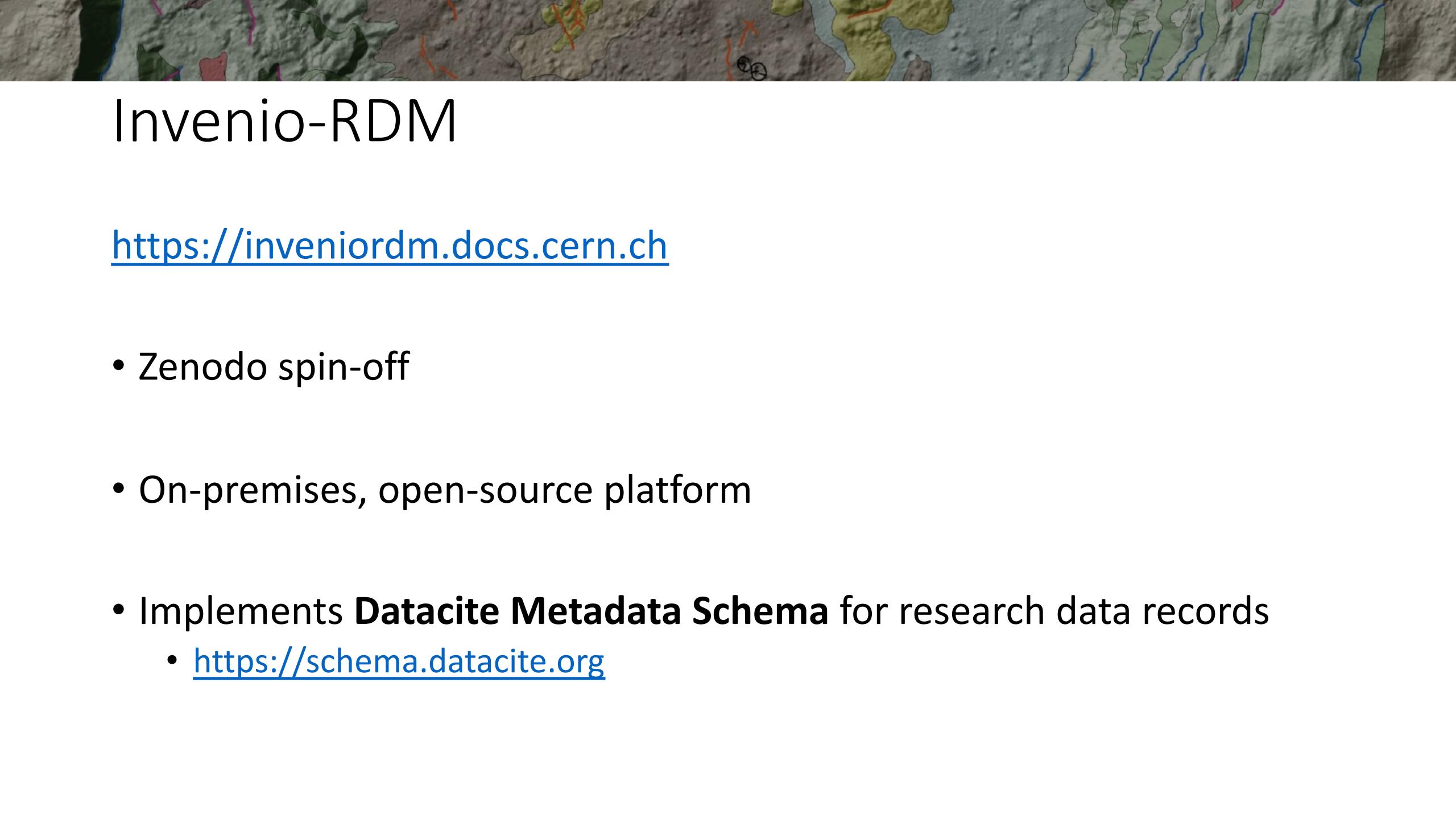


Research Data Management Systems

- Zenodo
 - <https://zenodo.org>
- Dataverse
 - <https://dataverse.org>
- CKAN
 - <https://ckan.org>
- FigShare
 - <https://figshare.com>
- FAIRDOM
 - <https://fairdomhub.org>

Research Data Management Systems

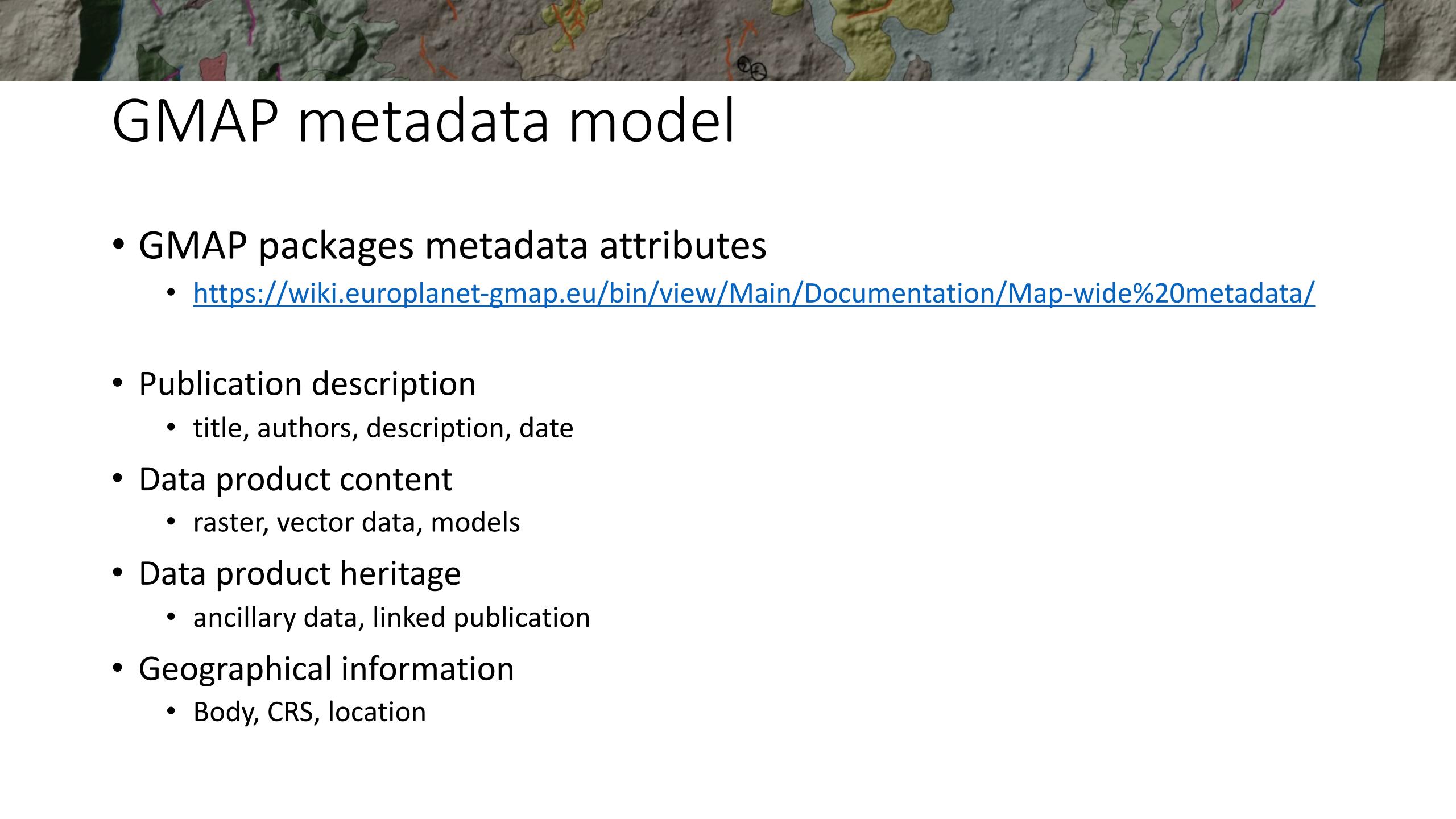
- Zenodo
 - <https://zenodo.org>
 - Dataverse
 - <https://dataverse.org>
 - CKAN
 - <https://ckan.org>
 - FigShare
 - <https://figshare.com>
 - FAIRDOM
 - <https://fairdomhub.org>
- 
- InvenioRDM**



Invenio-RDM

<https://inveniordm.docs.cern.ch>

- Zenodo spin-off
- On-premises, open-source platform
- Implements **Datacite Metadata Schema** for research data records
 - <https://schema.datacite.org>



GMAP metadata model

- GMAP packages metadata attributes
 - <https://wiki.europlanet-gmap.eu/bin/view/Main/Documentation/Map-wide%20metadata/>
- Publication description
 - title, authors, description, date
- Data product content
 - raster, vector data, models
- Data product heritage
 - ancillary data, linked publication
- Geographical information
 - Body, CRS, location



GMAP metadata model

- GMAP packages metadata attributes
- Publication description
 - title, authors, description, date
- Data product content
 - raster, vector data, models
- Data product heritage
 - ancillary data, linked publication
- Geographical information
 - Body, CRS, location

Attribute	Description
Acknowledgements	Free-text acknowledge
Authors*	Semi-colon separated list of authors
Bounding box - Max Lat	Maximum latitude in degrees (-90:90] (> Min Lat)
Bounding box - Max Lon	East-most Longitude in degrees (-180:180] (> Min Lon)
Bounding box - Min Lat	Minimum latitude in degrees [-90:90) (< Max Lat)
Bounding box - Min Lon	West-most Longitude in degrees [-180:180) (< Max Lon)
DOI of companion paper	DOI of linked publication
Data used	Semi-colon separated list of ancillary, original data used
Map name (GMAP_ID)*	Unique package name (GMAP-{{target-body}}-{{content-type}}-{{region-label}}_{{detail-label}})
Original Coordinate Reference System	WKT declaring map' CRS
Short description*	Free-text (500 words maximum) describing the map
Standards adhered to	Semi-colon list of standards used in the map
Stratigraphic info	Description of stratigraphic elements in the map
Target body*	Name of target body (eg, Mercury)
Title of map*	Map title (eg, Awesome Geologic Map of the region X)
Units Definition	Units color definition, polygon styling
?	
(?) Aims	Reason, goal for this map
(?) Heritage used	heritage information
(?) Link to other data	Links to extenal resources
(?) Other comments	free-text (notes, errata, warnings)
(?) Output scale	Map spatial scale
(?) Related products	Other geological maps complementing this one
(?) Type	Either "draft" or "released"
Files	
• document/*	Documents, raster and vector data. Mandatory: map in PDF format.
• raster/	
• vector/	

InvenioRDM metadata model

- Invenio is a general-purpose RDM system
 - Record metadata not meant for planetary data (*e.g.*, no location or CRS)

The screenshot displays the InvenioRDM record editor interface, divided into two main sections: 'Basic information' on the left and 'Recommended information' on the right.

Basic information:

- Files:** A section for managing uploaded files.
- Resource type:** A dropdown menu for selecting the type of resource.
- Title:** A text input field for the title of the record.
- Publication date:** A date input field set to "2023-06-21". A note below specifies the format as YYYY-MM-DD, YYYY-MM, or YYYY.
- Creators:** A section for adding creators, with a rich-text editor for their descriptions.
- Licenses:** A section for selecting a license, currently set to "Creative Commons Attribution 4.0 International". A note explains the terms of the license.

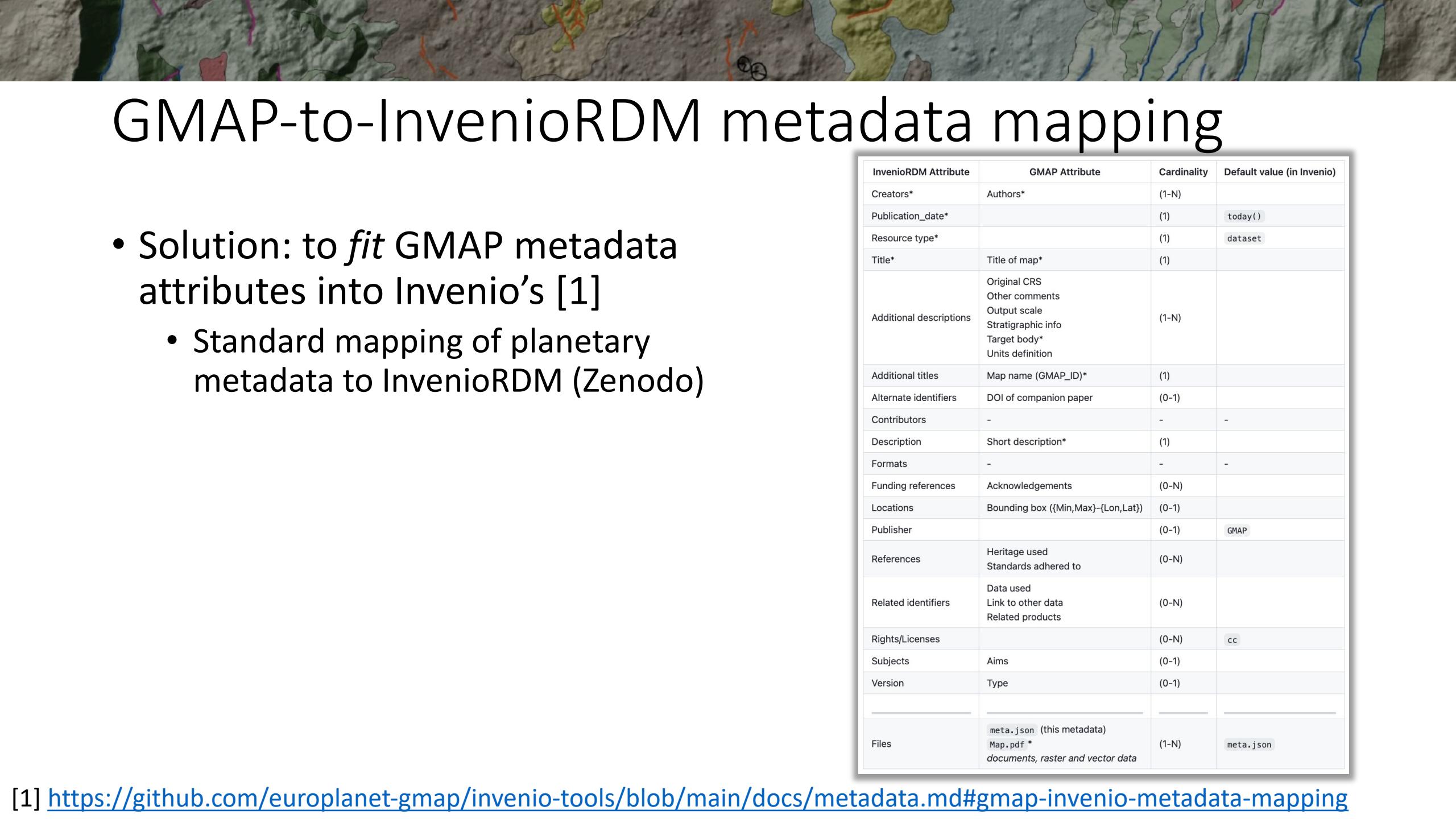
Draft: Buttons for "Save draft" and "Preview". A large green "Publish" button is prominently displayed.

Visibility: Options for "Full record" (Public or Restricted) and "Files only" (Public or Restricted). The "Public" option is selected for both.

Options: A section for applying an embargo.

Recommended information:

- Contributors:** A section for adding contributors.
- Subjects:** A section for adding subjects, with a search bar for "Search for a subject by name".
- Languages:** A section for adding languages, with a search bar for "Search for a language by name (e.g. "eng", "fr" or "Polish")".
- Dates:** A section for adding dates, with a note about the format: DATE or DATE/DATE where DATE is YYYY or YYYY-MM or YYYY-MM-DD.
- Version:** A section for adding version numbers.
- Publisher:** A section for adding the publisher's name, currently set to "GMAP".
- Funding:** A section for adding funding information.
- Alternate identifiers:** A section for adding alternate identifiers, with a "Related works" subsection.



GMAP-to-InvenioRDM metadata mapping

- Solution: to *fit* GMAP metadata attributes into Invenio's [1]
 - Standard mapping of planetary metadata to InvenioRDM (Zenodo)

InvenioRDM Attribute	GMAP Attribute	Cardinality	Default value (in Invenio)
Creators*	Authors*	(1-N)	
Publication_date*		(1)	today()
Resource type*		(1)	dataset
Title*	Title of map*	(1)	
Additional descriptions	Original CRS Other comments Output scale Stratigraphic info Target body* Units definition	(1-N)	
Additional titles	Map name (GMAP_ID)*	(1)	
Alternate identifiers	DOI of companion paper	(0-1)	
Contributors	-	-	-
Description	Short description*	(1)	
Formats	-	-	-
Funding references	Acknowledgements	(0-N)	
Locations	Bounding box ({Min,Max}-{Lon,Lat})	(0-1)	
Publisher		(0-1)	GMAP
References	Heritage used Standards adhered to	(0-N)	
Related identifiers	Data used Link to other data Related products	(0-N)	
Rights/Licenses		(0-N)	cc
Subjects	Aims	(0-1)	
Version	Type	(0-1)	
Files	meta.json (this metadata) Map.pdf* documents, raster and vector data	(1-N)	meta.json

[1] <https://github.com/europlanet-gmap/invenio-tools/blob/main/docs/metadata.md#gmap-invenio-metadata-mapping>

GMAP-to-InvenioRDM metadata mapping

- Solution: to *fit* GMAP metadata attributes into Invenio's [1]
 - Standard mapping of planetary metadata to InvenioRDM (Zenodo)

InvenioRDM Attribute	GMAP Attribute	Cardinality	Default value (in Invenio)
Creators*	Authors*	(1-N)	
Publication_date*		(1)	today()
Resource type*		(1)	dataset
Title*	Title of map*	(1)	
Additional descriptions	Original CRS Other comments Output scale Stratigraphic info Target body* Units definition	(1-N)	
Additional titles	Map name (GMAP_ID)*	(1)	
Alternate identifiers	DOI of companion paper	(0-1)	
Contributors	-	-	-
Description	Short description*	(1)	
Formats	-	-	-
Funding references	Acknowledgements	(0-N)	
Locations	Bounding box ({Min,Max}-{Lon,Lat})	(0-1)	
Publisher		(0-1)	GMAP
References	Heritage used Standards adhered to	(0-N)	
Related identifiers	Data used Link to other data Related products	(0-N)	
Rights/Licenses		(0-N)	cc
Subjects	Aims	(0-1)	
Version	Type	(0-1)	
Files	meta.json (this metadata) Map.pdf* documents, raster and vector data	(1-N)	meta.json

[1] <https://github.com/europlanet-gmap/invenio-tools/blob/main/docs/metadata.md#gmap-invenio-metadata-mapping>

Examples

- <https://data.europlanet-gmap.eu/records/a8493-pz977>

Geomorphologica map of Sputnik Planitia, Pluto

Suarez Valencia, Javier Eduardo¹; Ochoa Gutierrez, Luis Hernan²; Delgado-Correal, Camilo³; Saavedra Daza, Fabian²

Versions
Version v1 Feb 3, 2022

Details
Resource type: Figure
Publisher: GMAP

Rights
Creative Commons Attribution 4.0 International

Export
JSON Export

Citation

Style APA

Suarez Valencia, J. E., Ochoa Gutierrez, L. H., Delgado-Correal, C., & Saavedra Daza, F. (2022). Geomorphologica map of Sputnik Planitia, Pluto. GMAP.

Description

Geomorphological map of Sputnik Planitia and its surroundings. This is a young and geological active zone in Pluto, which is the result of a convective process of ice happening inside the basin.

Spatial Information:

- Bounding-Box: minlat = -25, maxlat = 50, westlon = -32, eastlon = 22
- Target body: Pluto

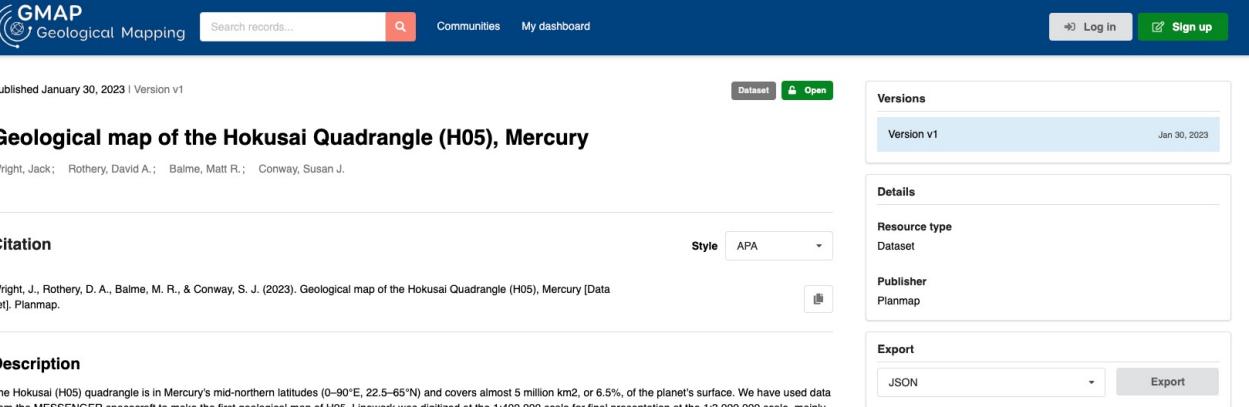
Ancillary information:

- Type: Released
- Output scale: 1:35.000.000
- Original Coordinate Reference System: Pluto_2000 (R: 1195000) Datum: D_pluto_2000
- Data used: LORRI (75m/p-250mp)
- Standards: Planetary mapping standards document
- Aims: Geomorphological mapping
- Map version: 1.0

Files

GeomorphologicalUnits.jpg

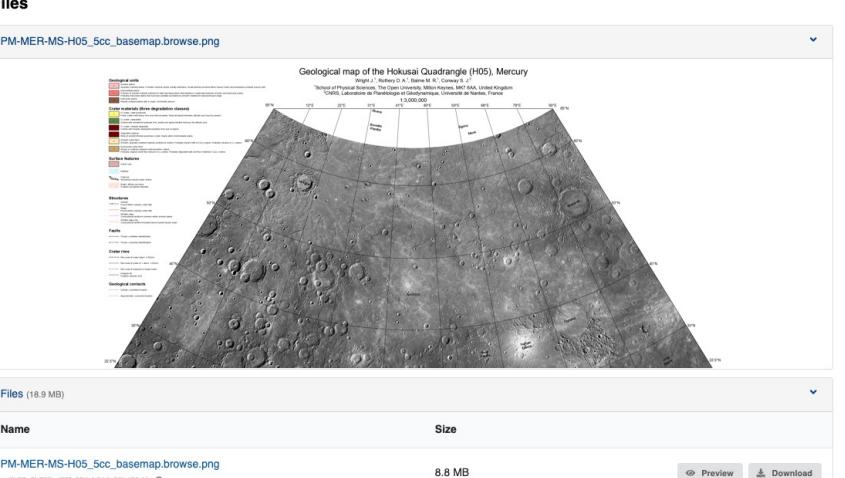
Files (16.7 MB)		
Name	Size	
GeomorphologicalUnits.jpg	7.5 MB	Preview Download
Ages.jpg	150.4 kB	Preview Download
Model.jpg	313.4 kB	Preview Download
Terrains.jpg	8.8 MB	Preview Download



The screenshot shows the GMAP Geological Mapping interface. At the top, there's a search bar with 'Search records...', a magnifying glass icon, and links for 'Communities' and 'My dashboard'. On the right, there are 'Log in' and 'Sign up' buttons. Below the header, a message says 'Published January 30, 2023 | Version v1'. There are 'Dataset' and 'Open' buttons. To the right, there's a 'Versions' section showing 'Version v1' from 'Jan 30, 2023'. A 'Details' section includes 'Resource type: Dataset', 'Publisher: Planmap', and an 'Export' button for 'JSON'. The main content area is titled 'Geological map of the Hokusai Quadrangle (H05), Mercury' by Wright, Jack; Rothery, David A.; Balme, Matt R.; Conway, Susan J. It features tabs for 'Citation' (with a style dropdown for APA) and 'Description'. The 'Description' tab contains detailed information about the map's creation, data sources, and classification schemes. It also lists 'Spatial information', 'Ancillary information', and 'Files'.

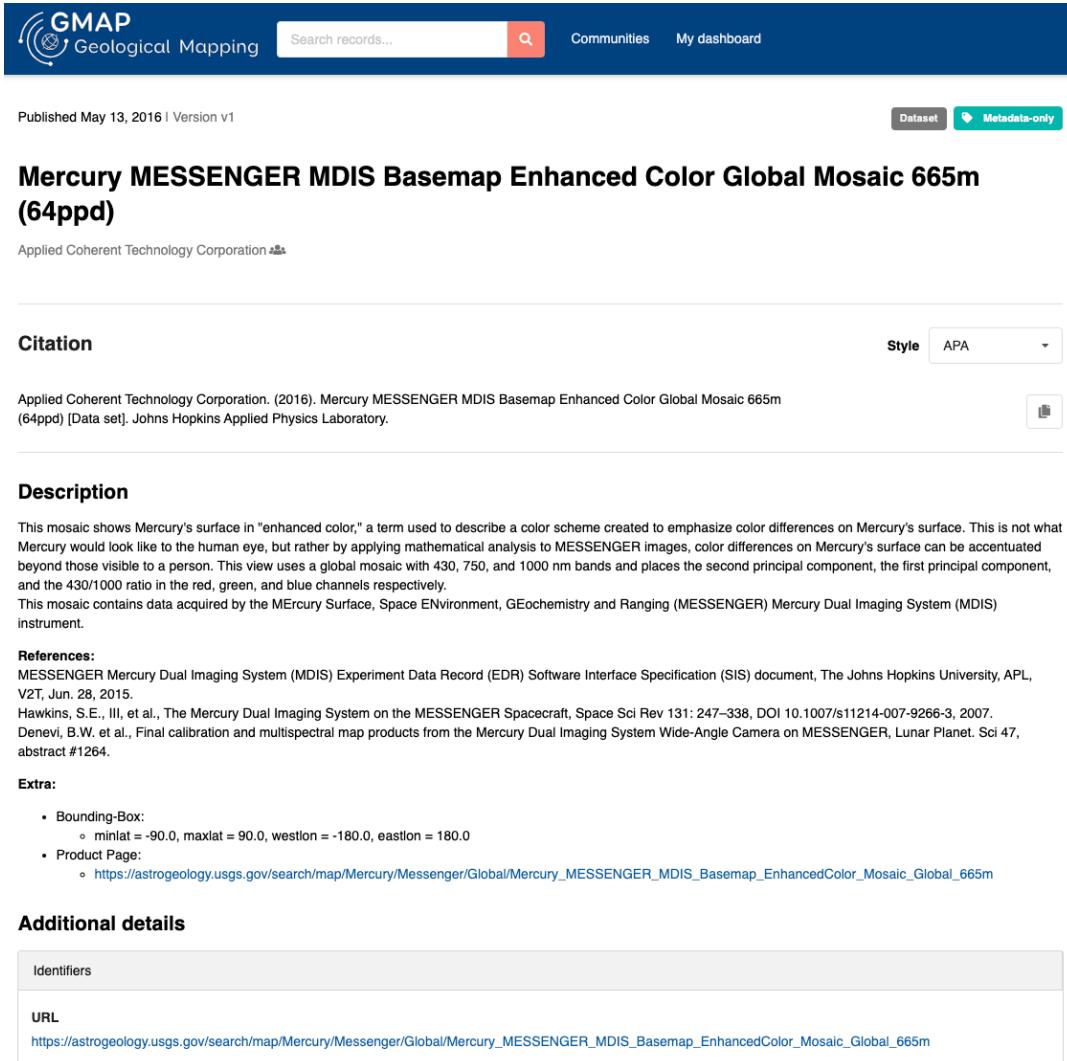
Examples

- <https://data.europlanet-gmap.eu/records/sqt7q-48y61>

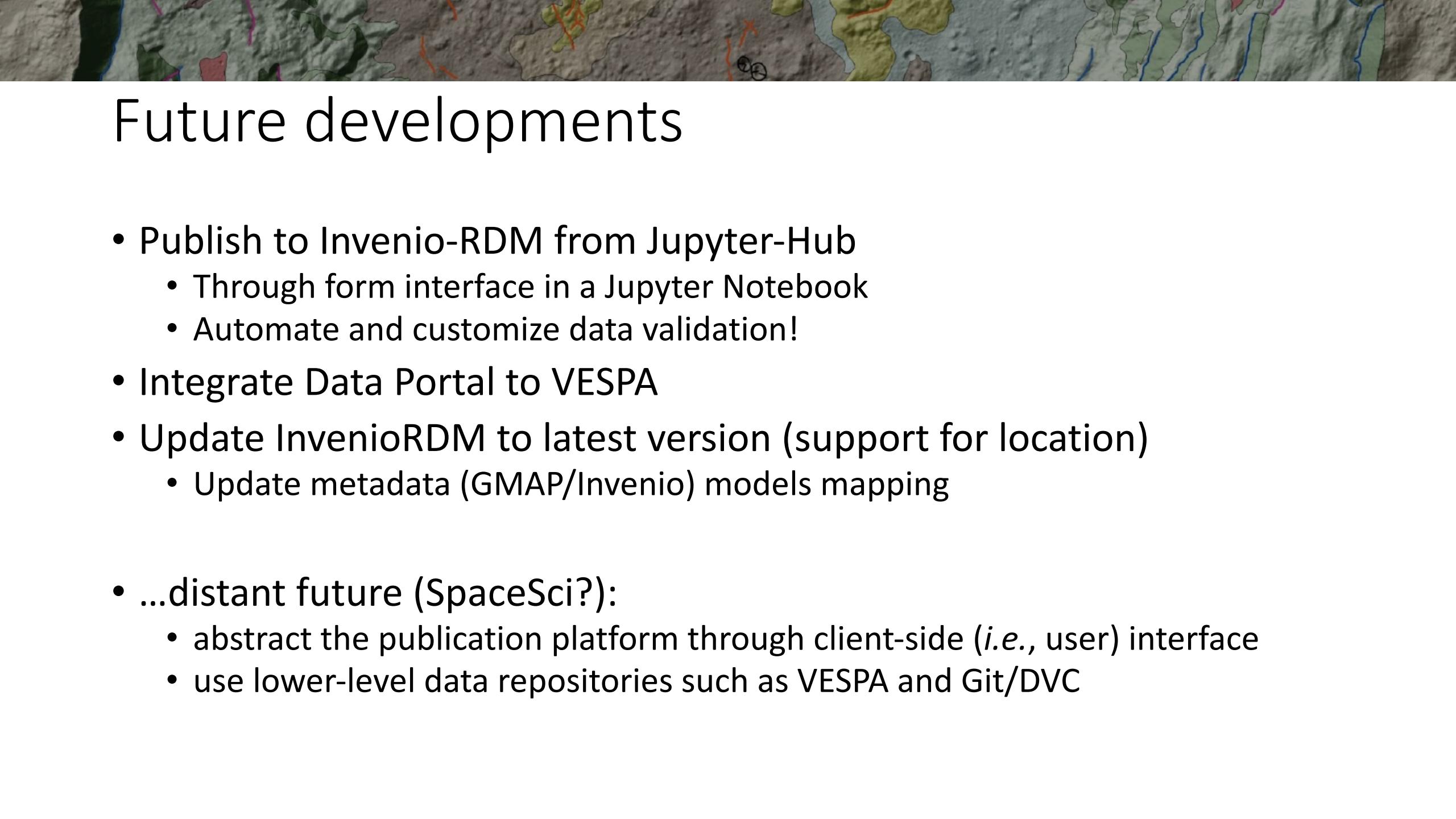


Planetary data index

- We use Invenio “metadata-only” records for “meta-publications”
 - Records that don’t host/hold the data
 - Point to an external archive (*e.g.*, Astropedia)
- <https://data.europlanet-gmap.eu/records/mxzdd-gds58>



The screenshot shows a dataset record on the GMAP Geological Mapping platform. The header includes the GMAP logo, a search bar, and links for Communities and My dashboard. Below the header, it says "Published May 13, 2016 | Version v1". A "Dataset" button and a "Metadata-only" badge are visible. The main title is "Mercury MESSENGER MDIS Basemap Enhanced Color Global Mosaic 665m (64ppd)". It is attributed to "Applied Coherent Technology Corporation". The "Citation" section uses APA style and provides the source information. The "Description" section details the mosaic's characteristics, mentioning "enhanced color" and its technical parameters. The "References" section lists the MESSENGER MDIS Experiment Data Record (EDR) Software Interface Specification (SIS) document and a paper by Hawkins et al. The "Extra" section contains a bounding box and a product page link. The "Additional details" section includes tabs for Identifiers and URL, with the URL being https://astrogeology.usgs.gov/search/map/Mercury/Messenger/Global/Mercury_MESSENGER_MDIS_Basemap_EnhancedColor_Mosaic_Global_665m.



Future developments

- Publish to Invenio-RDM from Jupyter-Hub
 - Through form interface in a Jupyter Notebook
 - Automate and customize data validation!
- Integrate Data Portal to VESPA
- Update InvenioRDM to latest version (support for location)
 - Update metadata (GMAP/Invenio) models mapping
- ...distant future (SpaceSci?):
 - abstract the publication platform through client-side (*i.e.*, user) interface
 - use lower-level data repositories such as VESPA and Git/DVC