

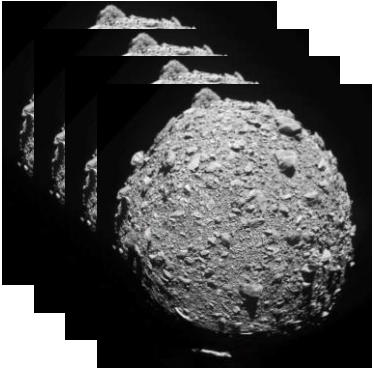
PROVEX – APPROACHABLE DATA PROVENANCE TRACKING FOR SPACE MISSIONS

Harald Steinlechner¹, Sophie Pichler¹, Tomas Khout^{2,4}, David Korda³

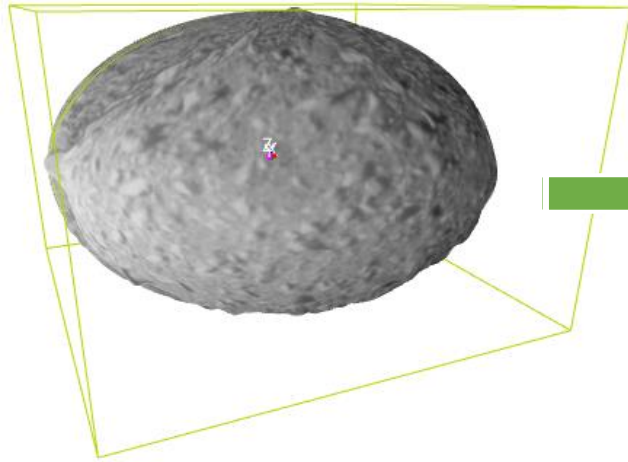
Thomas Ortner, Christoph Traxler¹, Gerhard Paar²



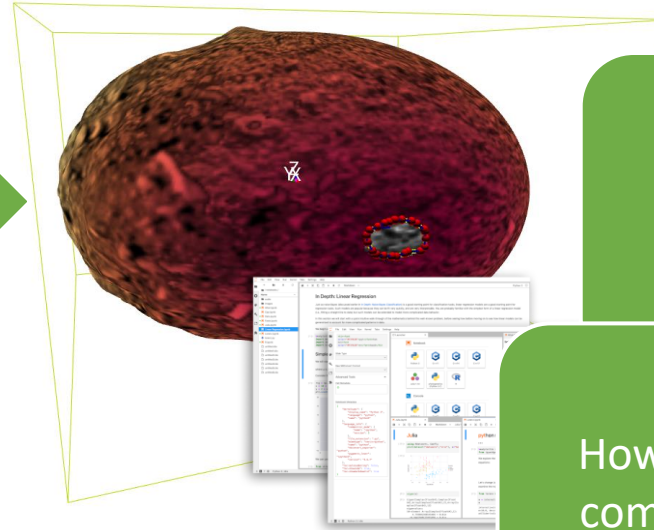
Image Data



3D Data



3D & data analysis



What data lead to what conclusions?

New reconstruction available, does it hold?

How does a new method compare to other results

What is the accuracy? Where does uncertainty come from?

Isn't there a principled approach?

Idea: Track provenance information throughout the whole data pipeline / scientific investigation

Provenance to the rescue

- Collaborative workflow management: “Provenance”
 - Provenance = **chronology of a scientific investigation**
 - Track provenance of all products of a complex, multidisciplinary scientific investigation
 - Answers the following questions:
 - **Who** was involved in the creation?
 - Which **methods** were applied?
 - Which **uncertainties** were added by these methods?
 - From which other products it was derived?
- **PROVEX** (Provenance Aware Planetary Science / Exploration Workflows)
 - Document entire history of scientific investigation

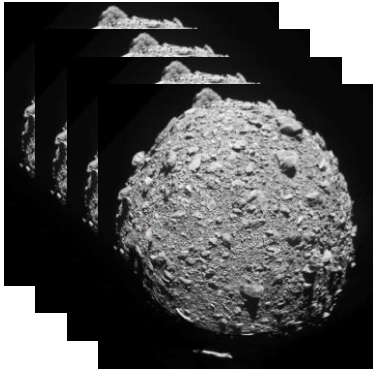
Content

- Motivation, Scope & Data
- Provenance and 3D data
- Applications & Showcases
- Integration with science workflows (notebooks)

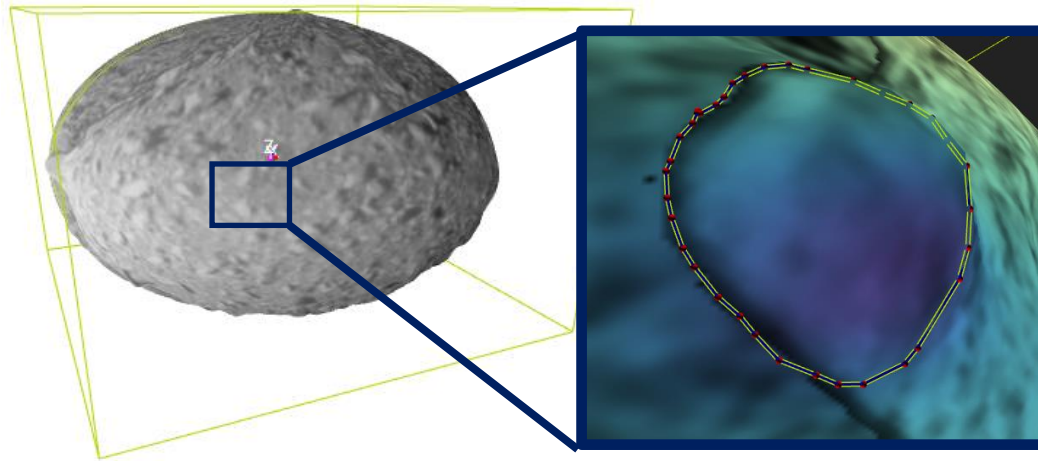
Scope & Data

- HERA as focused application case
- Brings together multiple tools

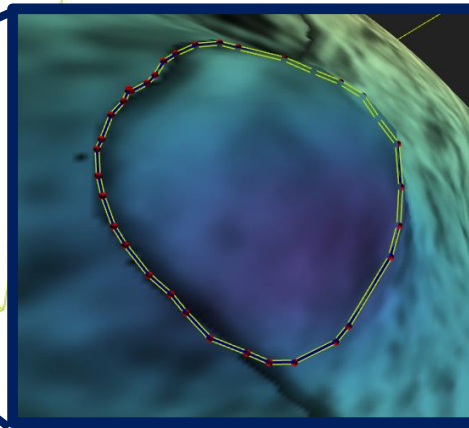
Image Data



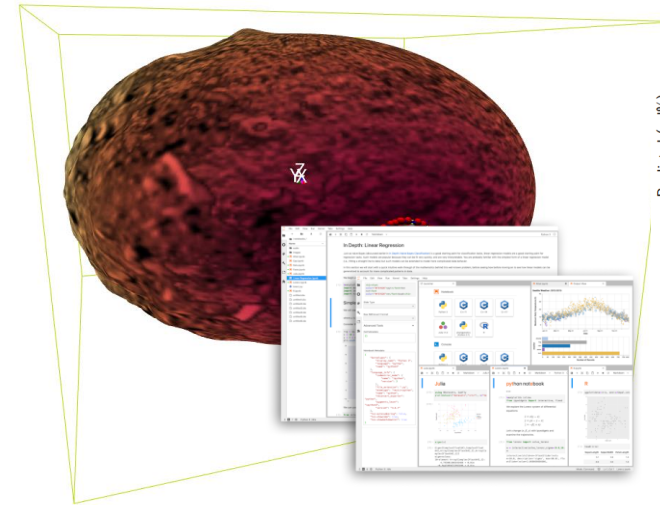
3D Data



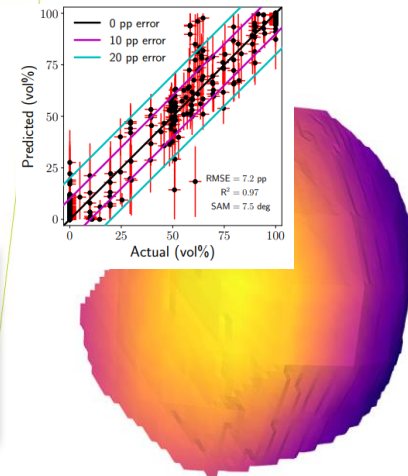
3D analysis



Scripts & notebooks



Presentations, papers etc



PROVEX components for HERA

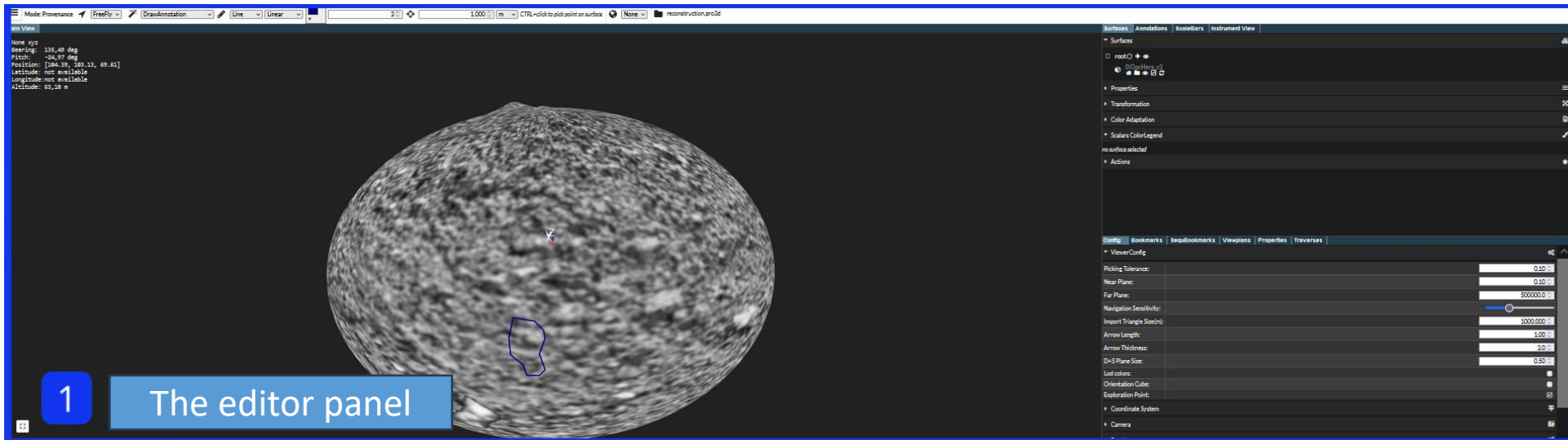
- **PRoViP**: Data & Data Product's Interface
 - Batch 3D Vision Processing
- **Data ingestion** pipeline
 - Looks at reconstruction meta-data and stores its provenance graph into graph database.
- **Graph database** which stores provenance information and metadata
- **PRo3D**: Realtime Rendering / Analysis
 - Huge multiscale 3D data
 - 3D GIS
 - Supports scientific operations
- **PRo3D notebook interface** for complex science use-cases.

Managed by a single web app

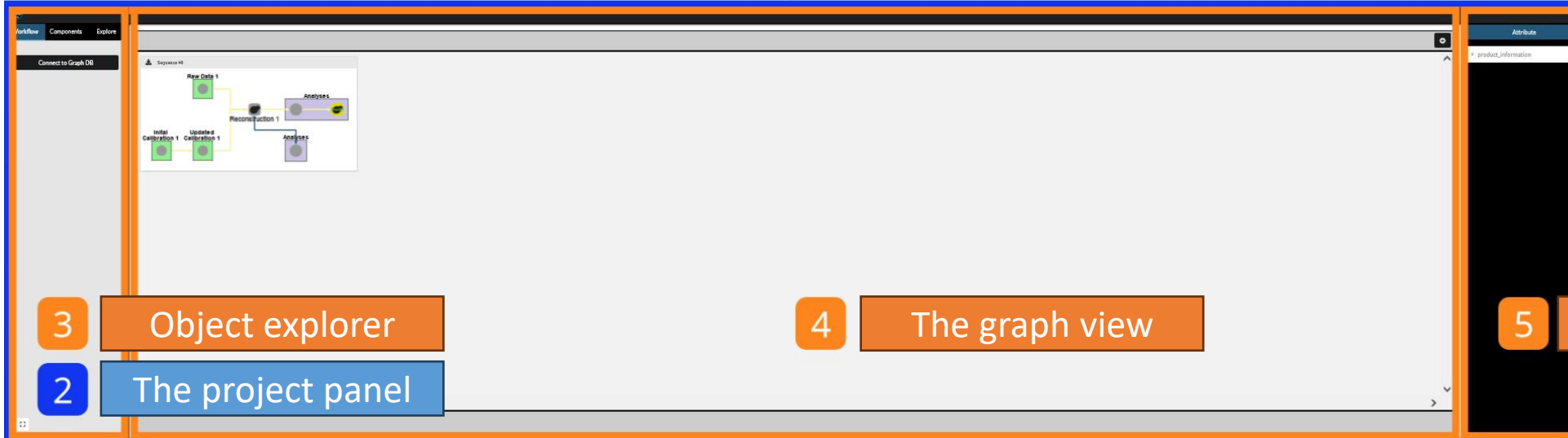
PRoViP

Scientific notebooks

Application Overview



1 The editor panel



3 Object explorer

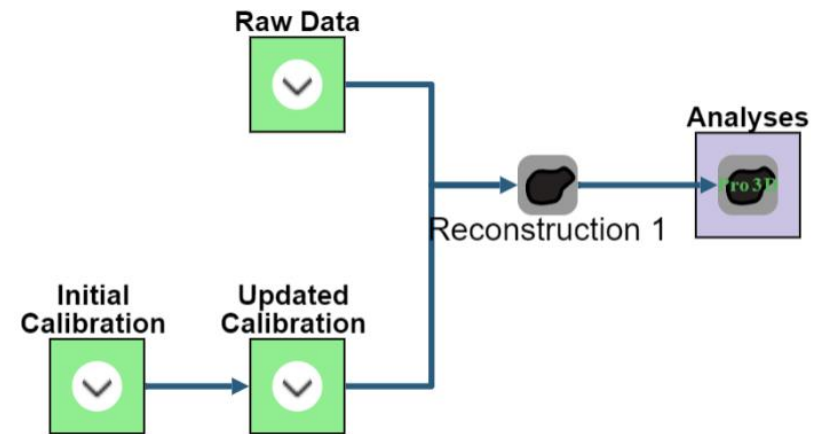
4 The graph view

5 Details & Attributes

2 The project panel

The PROVEX Provenance Graph

- Each involved product needs to
 - Know its sources of information
 - Involved uncertainty
 - Metadata such as author/investigator, date etc
- We store the data as a graph
 - Nodes: Products and derived products
 - Edges: Information flow & uncertainty

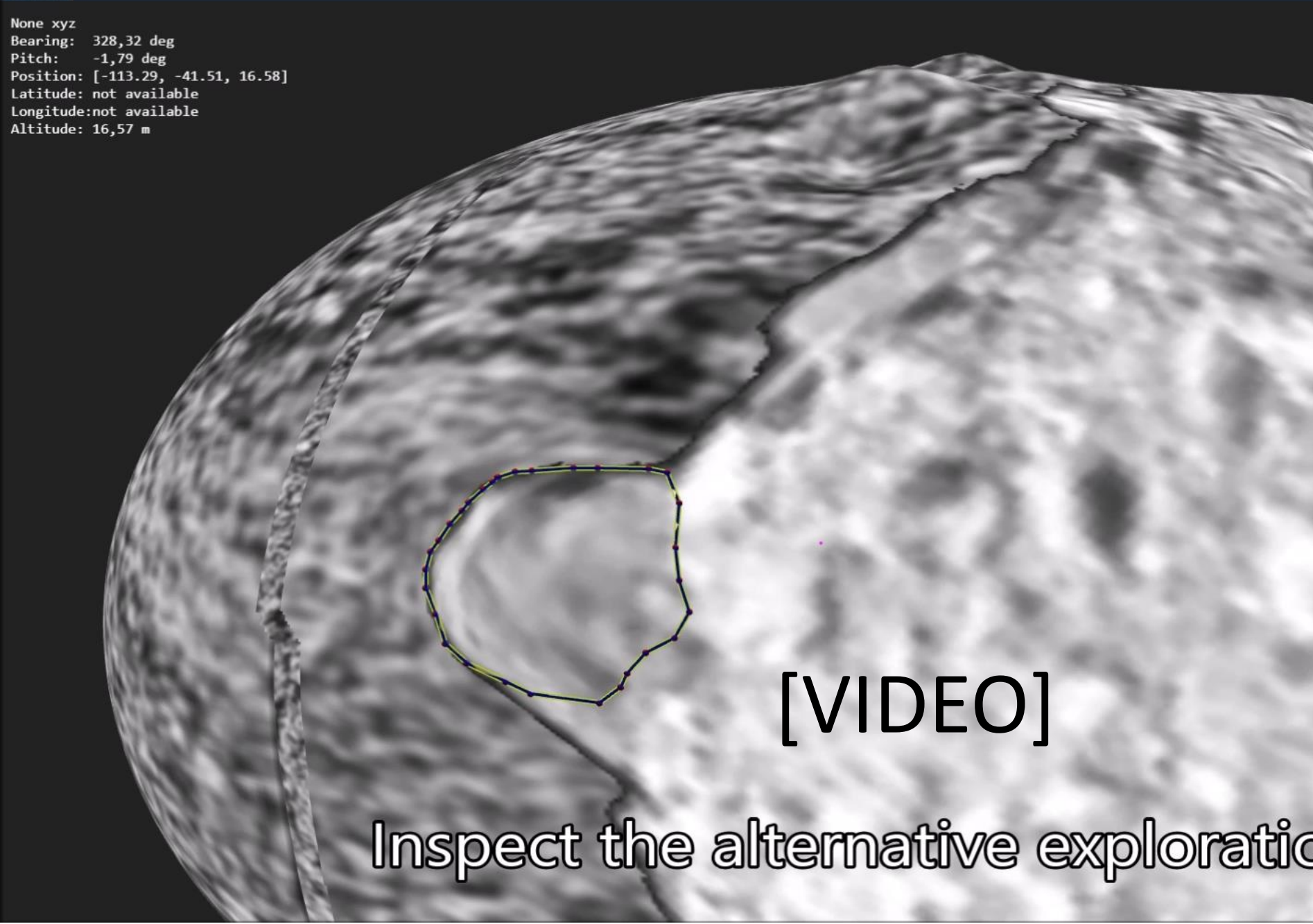


Types of data for provenance

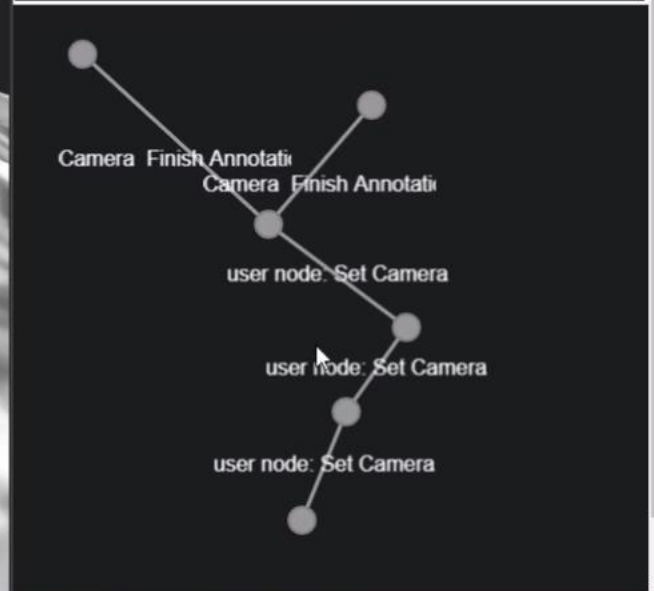
- Raw data
 - Instrument intrinsic and measurement parameters
- Ancillary data
 - SPICE orientations, instrument housekeeping – all “meta data”
- Processed data
 - Applied functions and algorithms such as transformations, reconstructions etc.
- Imported data
 - Conversion information
- Simulation result
 - Parameters of simulation runs
- Collaborative analysis result
 - Step-by-step documentation of analysis (for major steps)
 - Resulting presentations / papers / videos
- Input images
 - Instrument
 - Instrument calibration status
 - Sensing mission phase & time
 - Sensing geometry
 - Surface illumination status (global statistics)
 - (if available) capturing intent
- P_{Ro}ViP & Imported 3D Vision processing (data provenance)
 - History
 - Source data
 - Processing Parameters
 - Version, owner, Processing workflow used
 - Result statistics
- P_{Ro}3D visualization & exploration, annotations [screenshots] (data provenance)
 - Who (interaction provenance)
- P_{Ro}3D analysis
 - Bookmarks
 - Annotations
- P_{Ro}3D interpretation
 - (entities relevant for) Publications
- 3rd party presentation means
 - Power point slide presentations
 - YouTube video sequences
 - WinWord / LaTeX publication manuscripts

Provenance for reconstruction data

- When ingesting data into PROVEX, meta-data gets translated into the PROVEX provenance graph
- PRo3D, tracks provenance internally (fine-grained) and feeds it back into the PROVEX graph.



None xyz
Bearing: 328,32 deg
Pitch: -1,79 deg
Position: [-113.29, -41.51, 16.58]
Latitude: not available
Longitude: not available
Altitude: 16,57 m

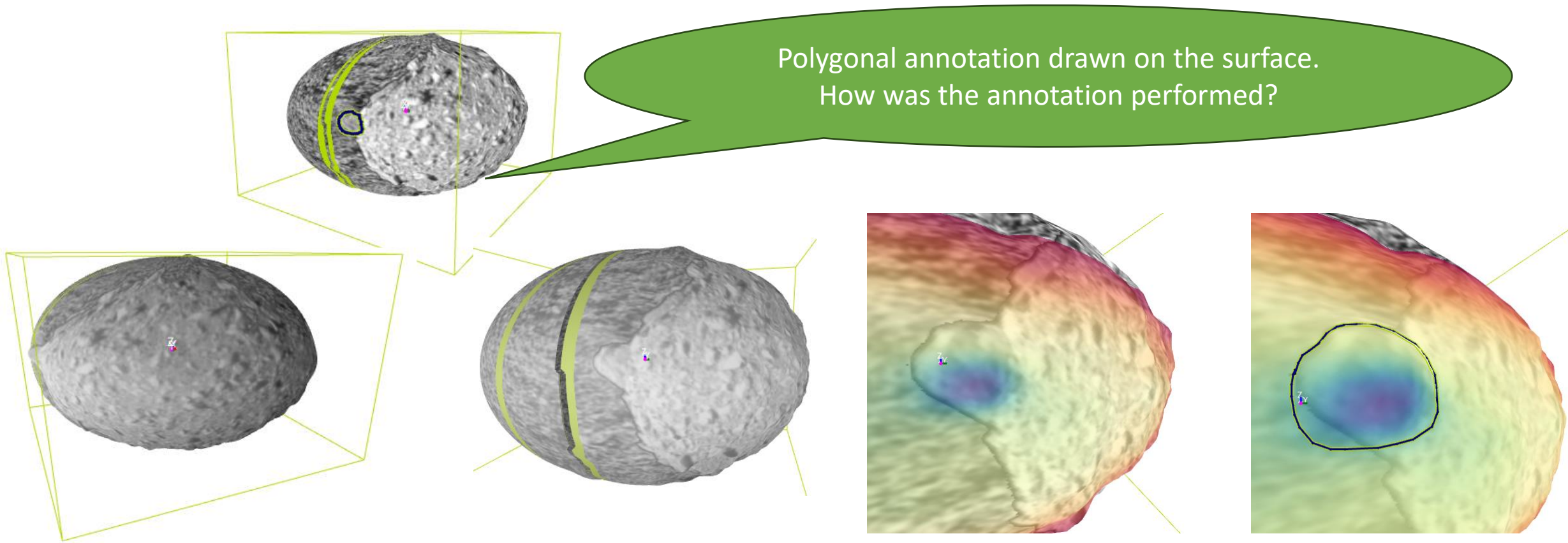


Exploration Point:	<input checked="" type="checkbox"/>
Coordinate System	
Pos:	[-0.74, -0.23, 0.96]
Up:	X: 0,000
	Y: 0,000
	Z: 1,000
North:	[1,00, 0,00, 0,00]
NorthO:	[1,00, 0,00, 0,00]
N-Offset:	0,0
Longitude:	NaN
Latitude:	NaN
Altitude:	NaN
Visible:	<input type="checkbox"/>
Location:	[-113.29, -41.51, 16.58]
Forward:	[0.851, 0.525, -0.031]

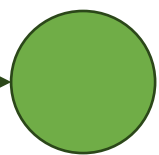
[VIDEO]

Inspect the alternative explorations.

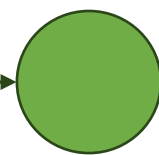
Example: Tracing back 3D operations



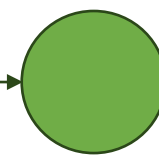
Data loaded



Camera adjusted



Render mode
adjusted



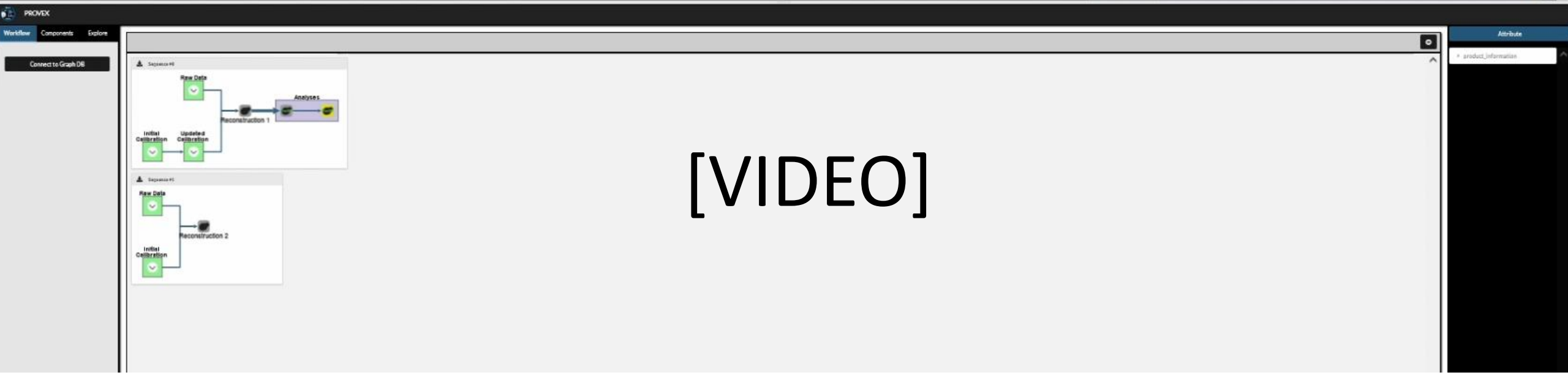
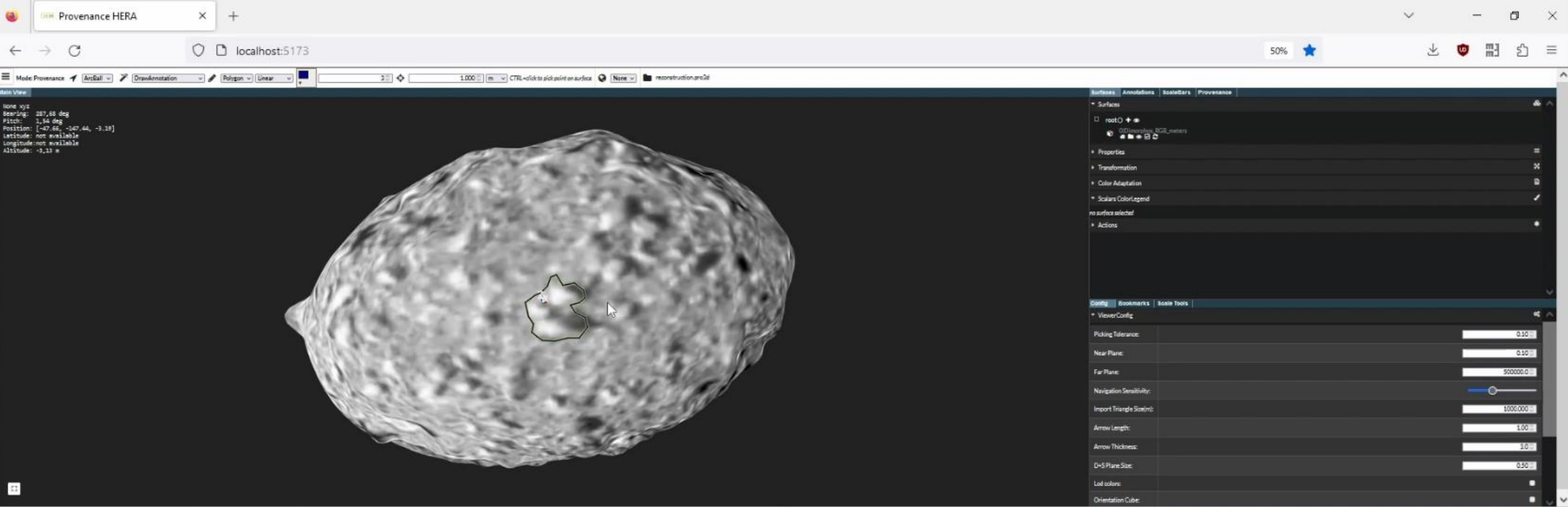
Created
measurement

Provenance for reconstruction data (2)

- Two modes...
- Manual tracking
 - important steps annotated by user
 - underlying provenance information tracked transparently to the user
- Fully automatic tracking
 - Heuristics for detecting „important“ steps in analysis
- Go back and forth in time at any time
- Information automatically stored in PROVEX database

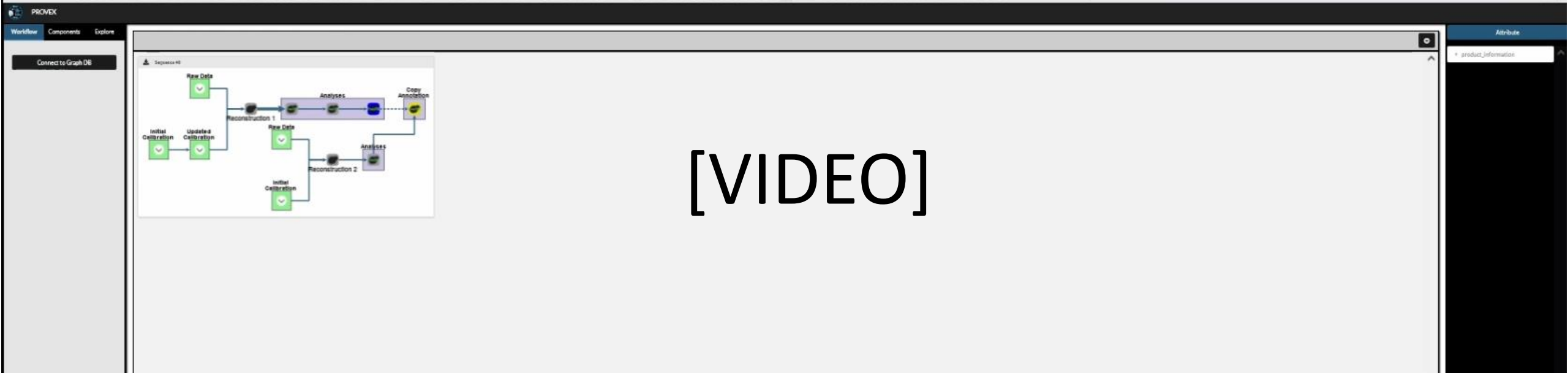
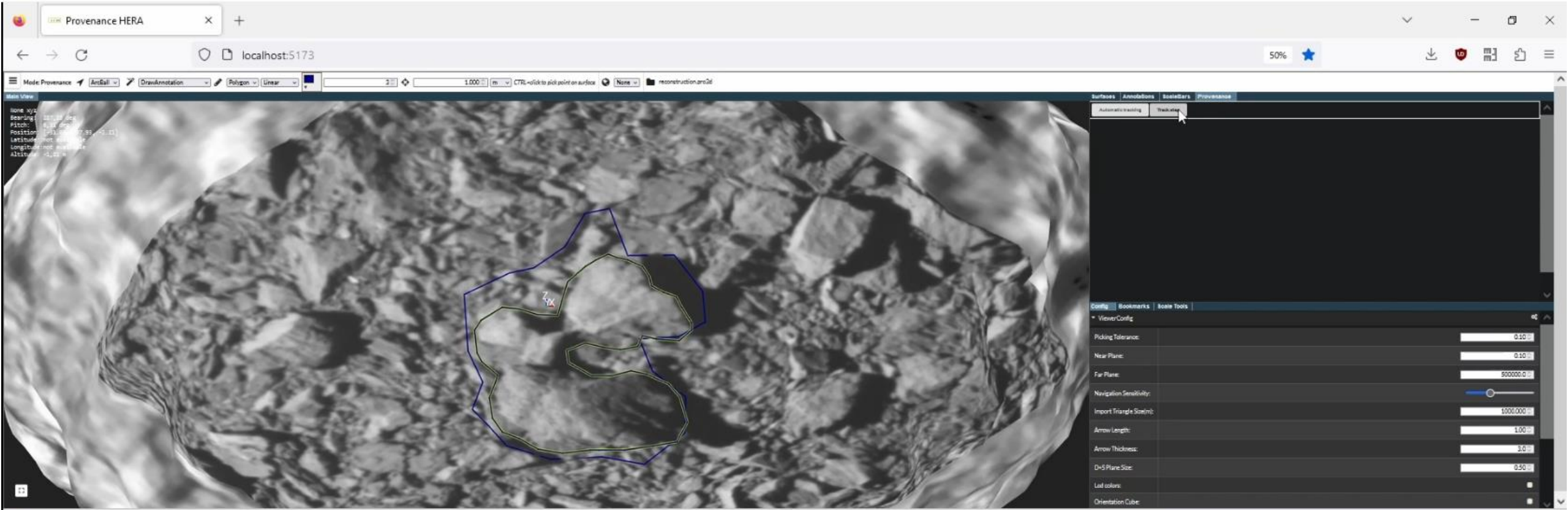
Working with 3D reconstructions

Fine-grained provenance tracking



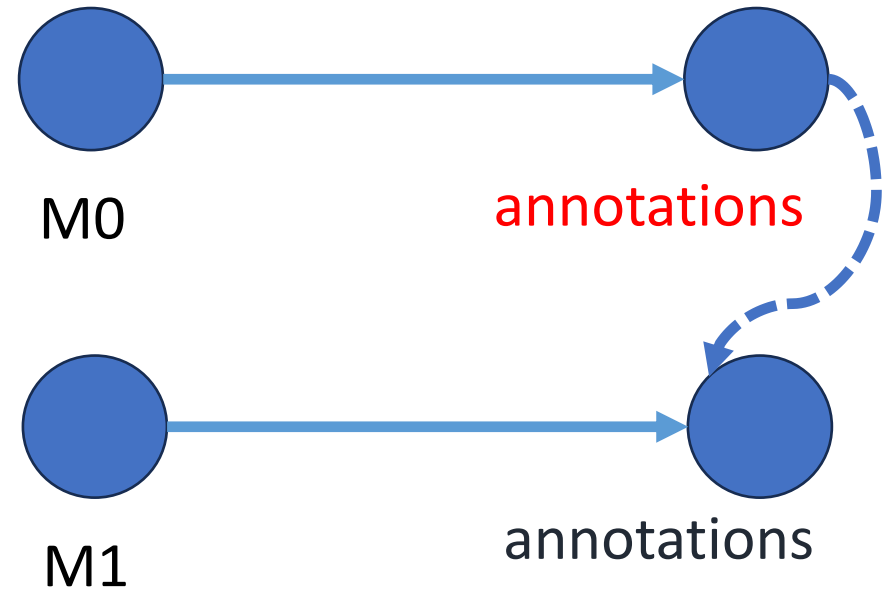
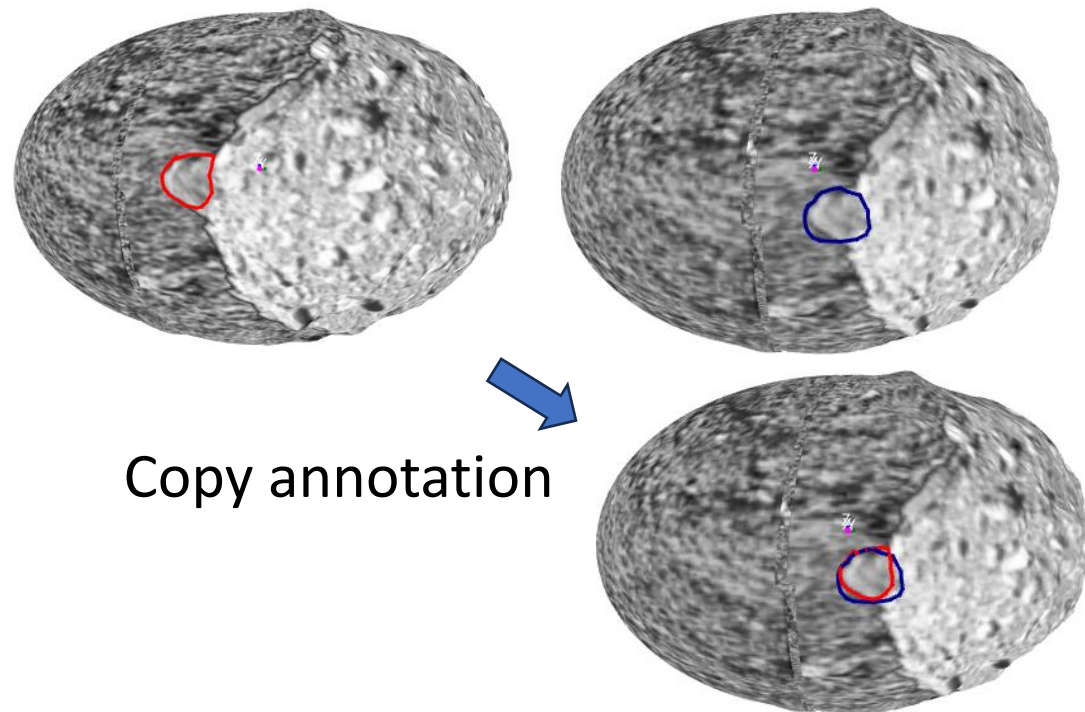
Next mission phase

Provenance tracking for multiple reconstructions



[VIDEO]

Provenance for multiple mission phases



Custom scientific evaluations

Interacting with PROVEX

Notebooks and science use cases

- Cannot cover all science use cases
- Science workflow = multidisciplinary
- Heterogenous set of tools used
- Get all of them integrated into PROVEX
 - PROVEX data API
 - Can be used from e.g. Jupyter Notebooks

```
File Edit Selection View Go Run Terminal Help
CutoutDemo.ipynb - proves - Visual Studio Code
CutoutDemo.ipynb •
CutoutDemo.ipynb > id = getAnnotation() # just use the first annotation for simplicity of the demo.
Code + Markdown | Run All | Clear Outputs of All Cells | Restart | Variables | Outline ... Python 3.7.0 64-bit
18 file.write(response.text)
19 file.close()
20
21 def getAnnotationGeometry(id, mappedLayers):
22     url = "http://localhost:4321/api/queries/queryAnnotationAsJson"
23     payload = {"annotationId": id, "queryAttributes": [mappedLayers],
24     response = requests.get(url, json=payload)
25     print(response.text)
Python
1 id = getAnnotation() # just use the first annotation for simplicity of
2
3 # instead of Ele.aara any other
4 # attribute available in the reconstruction can be used
5 print(getAnnotationGeometry(id, "Ele.aara"))
Python
Output exceeds the size limit. Open the full output data in a text editor
b0f7d700-73b2-40dc-a8b9-95cb7439287b
{
  "filteredPatches": [
    {
      "verticesWorldSpace": [
        [-86.12904442285547,
        -24.144038858862302,
        8.642272694420166
        ],
        [-86.0725563857461,
        -24.128204050512693,
        8.447665913414307
        ],
        [-86.30142296289453,
        -23.92193764731445,
        8.482631905388184
        ],
        [-86.12904442285547,
        -24.144038858862302,
        8.642272694420166
        ]
      ]
    }
  ]
}
Können Sie sich bitte eine Minute Zeit nehmen, um uns über
Ihre Nutzungserfahrung in VS Code zu informieren?
Source: Jupyter | Ja, nehmen Sie jetzt an der Umfrage teil | Nein, danke
```

VRo3D Viewer - 4.12.0-prepare9 - VRVis for Virtual Reality und Visualisierung Forschungs-GmbH

Mode: Provenance | FreeFly | DrawAnnotation | Polygon | Linear | 3

Main View

None xyz
Bearing: 357,19 deg
Pitch: -1,78 deg
Position: [-251.42, -26.76, 10.03]
Latitude: not available
Longitude: not available
Altitude: 10,02 m

Surfaces Annotations ScaleBars Provenance

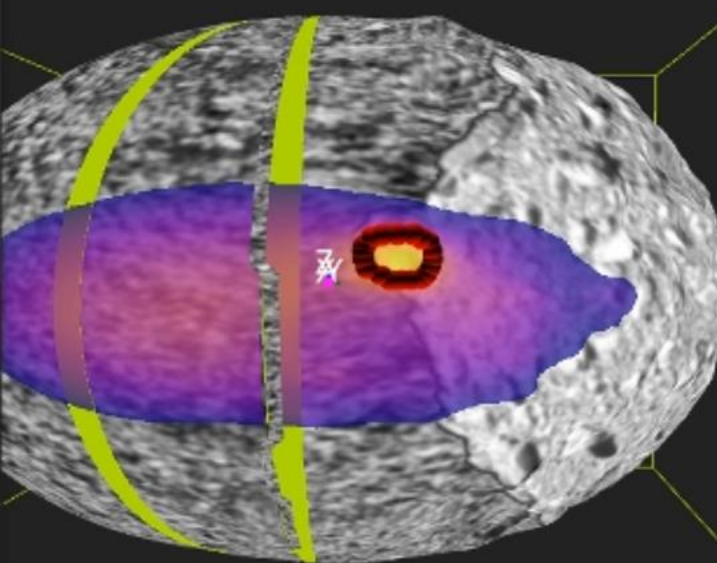
Fillmode: Fill
Scalers: None
Cull Faces: None
Set Homeposition:
OPCx Info path: none
Primary Texture: Texture
Transfer Function: Ramp
Secondary Textures: Ele
Texture Combiner: Blend
Blend Factor:
Min: 0.77
Max: 1
Color Map: plasma

Transformation
Color Adaptation

Config Bookmarks Scale Tools

Viewer Config

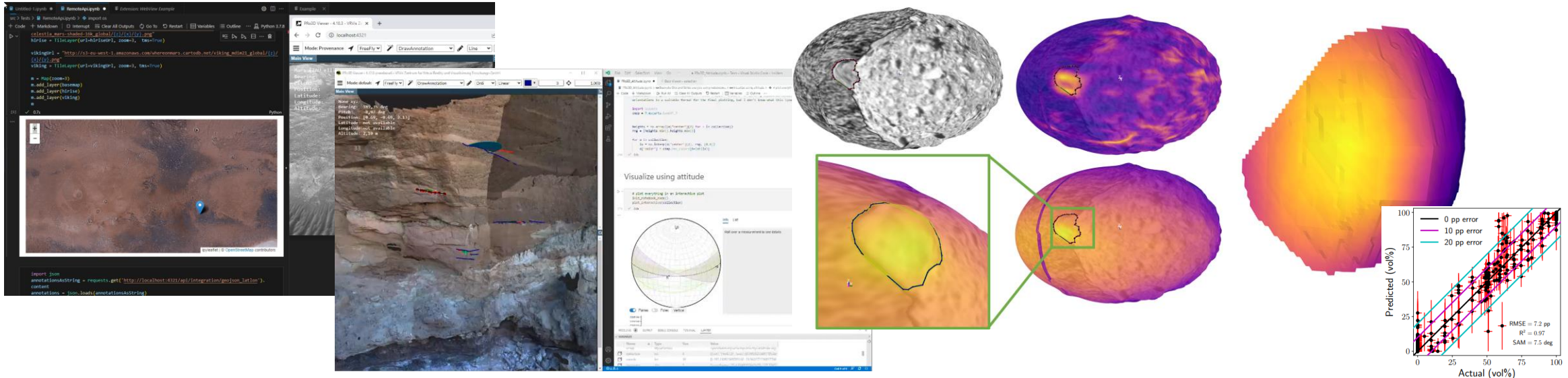
Picking Tolerances: 0.10
Near Plane: 0.10
Far Plane: 500000.0
Navigation Sensitivity:
Import Triangle Size(m): 1000.000
Arrow Length: 1.00
Arrow Thickness: 3.0
D+S Plane Size: 0.50
Lod colors:
Orientation Cube:
Exploration Point:
Coordinate System:
Camera:
Frustum:



[VIDEO]

Advantages

- Easier to comprehend scientific workflows
- Reproducible and traceable results
- Hub for collaborative activity
- Tracking uncertainties increases confidence in results
- Efficiently (re)create results for presentation and dissemination



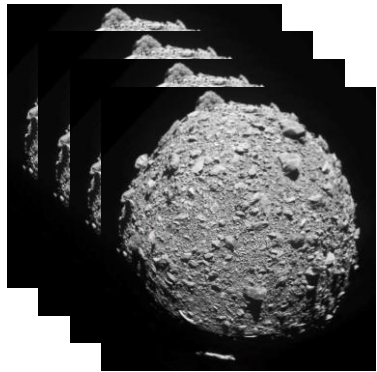
Future work & activities

- Exploration of additional science use cases
- Integration of broad set of tools

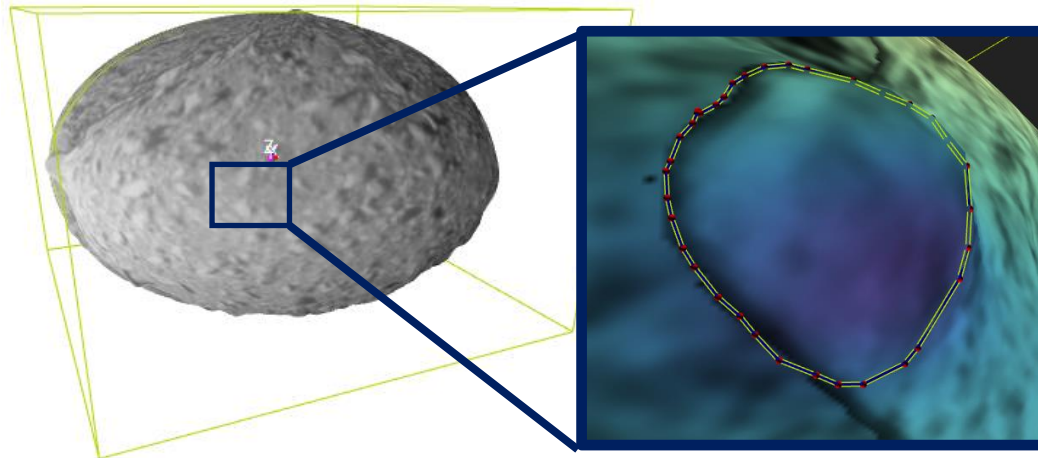
What about my workflow?

What about my science tools?

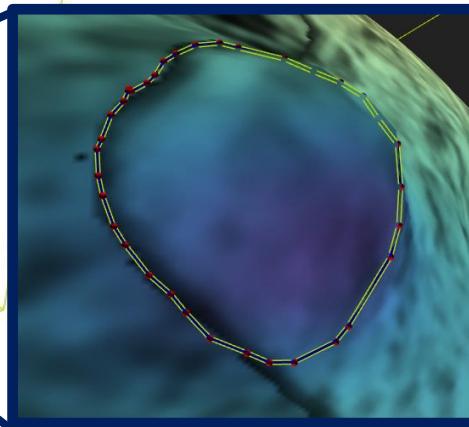
Image Data



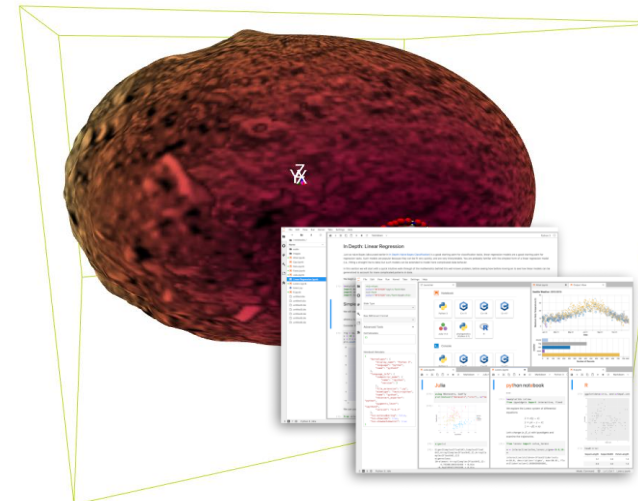
3D Data



3D analysis



Scripts & notebooks



Presentations, papers etc

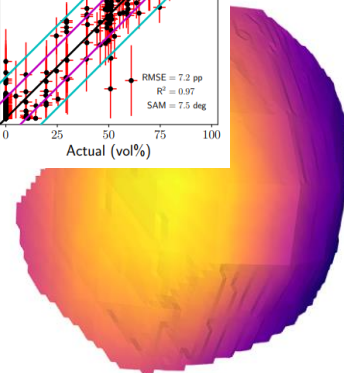
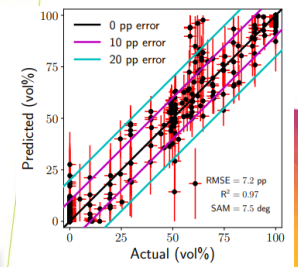
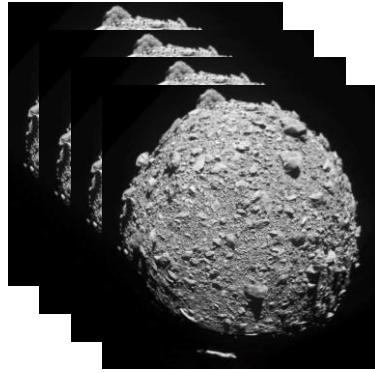
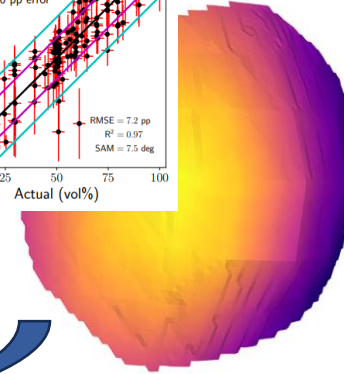
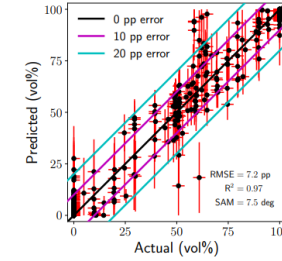
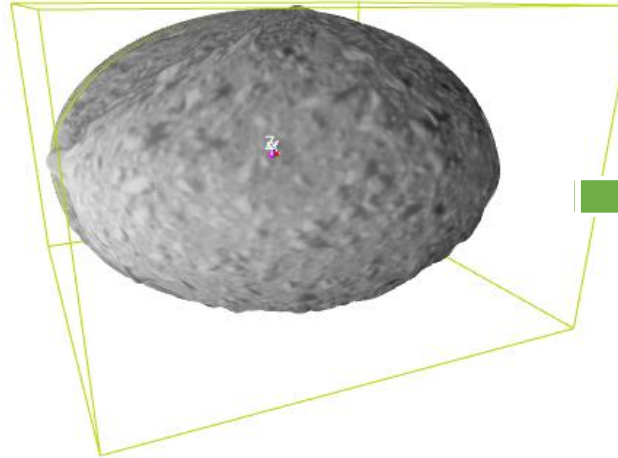






Image Data



3D Data



-  Q & A
-  Let's try out our scientific workflow
-  Request for contributions
-  <https://pro3d.space>, <https://github.com/pro3d-space/>

