Ten Years of OPUS: Lessons Learned

Mark Showalter, Rob French, Mitch Gordon, Matt Tiscareno, Mike Evans

Three verbs…

- **“Search”**: to find data that you think probably exists.
- **“Discover”**: to find data that you want but didn’t know exists.
- **“Explore”**: to find data that exists but you didn’t know you want.

PDS needs to support all three!

What do we need?

- For **Search**:
  - Reliable, surprise-free metadata
- For **Discovery**:
  - Consistent metadata across missions and instruments
  - A hierarchical database schema
- For **Exploration**:
  - Detailed metadata
  - Sensible granularity
  - Informative browse products

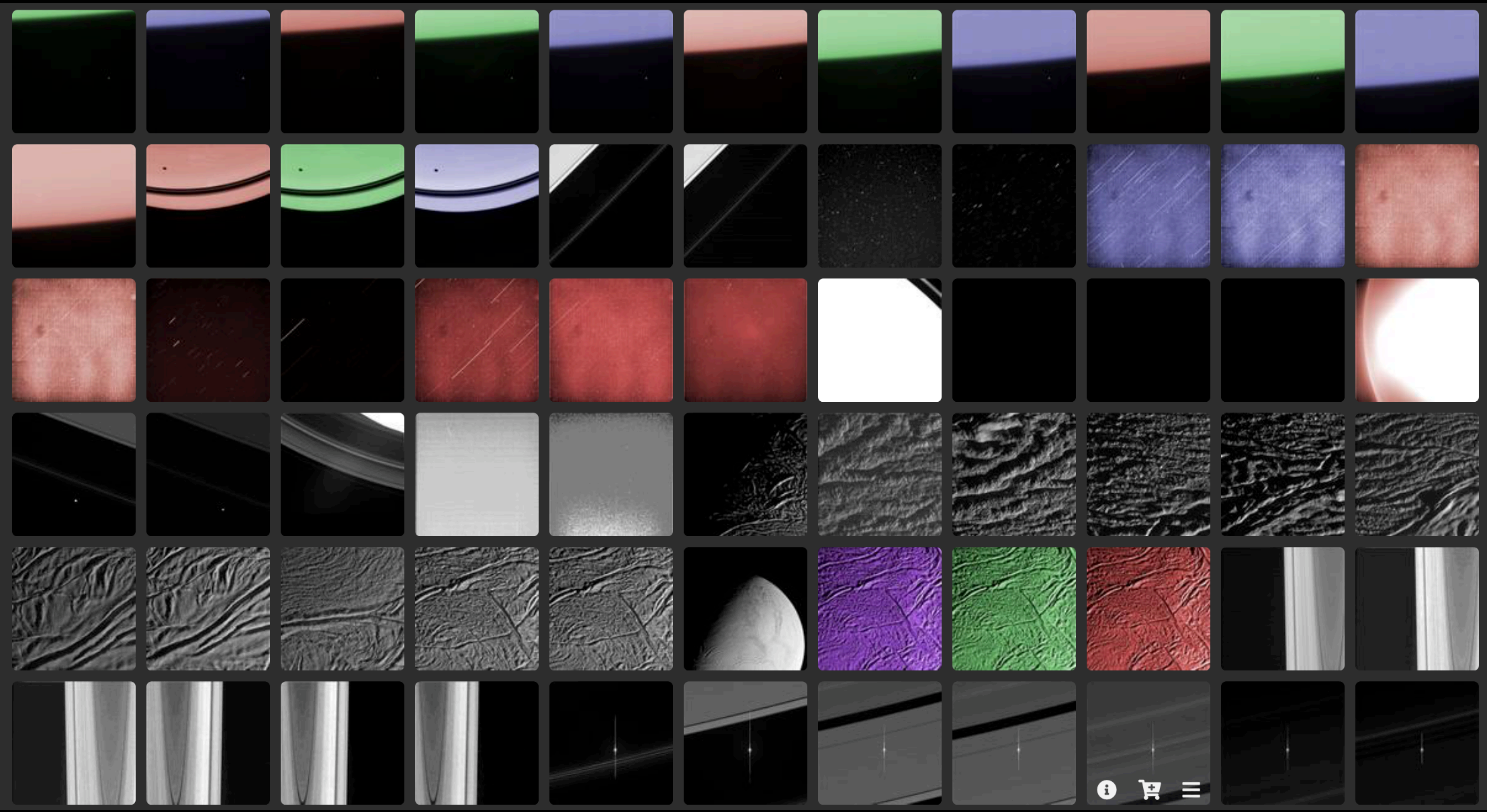
What do we need to do to support search?

(Finding data that you think probably exists)

Needed: Reliable Metadata

- What's do users expect from our metadata?
  - No surprises.
  - No need for "secret knowledge".

Cassini ISS, Target Name = “Sky”



Needed: Reliable Metadata

Every Cassini ISS target name since 2004

- AEGAEON
- ALBIORIX
- ANTHE
- ATLAS
- BEBHIONN
- BERGELMIR
- BESTLA
- CALYPSO
- DAPHNIS
- DIONE
- EARTH
- ENCELADUS
- EPIMETHEUS
- **ERRIAPO**
- ERRIAPUS
- FORNJOT
- GREIP
- HATI
- HELENE
- HYPERION
- **HYROKKIN**
- IAPETUS
- IJIRAQ
- JANUS
- JARNSAXA
- JUPITER
- **K07S4**
- KARI
- KIVIUQ
- LOGE
- METHONE
- MIMAS
- MUNDILFARI
- NARVI
- PAALIAQ
- PALLENE
- PAN
- PANDORA
- PHOEBE
- POLYDEUCES
- PROMETHEUS
- RHEA
- **S12_2004**
- **S13_2004**
- **S14_2004**
- **S18_2004**
- **S8_2004**
- SATURN
- SIARNAQ
- **SKADI**
- SKATHI
- SKOLL
- **SKY**
- SUN
- SURTUR
- **SUTTUNG**
- SUTTUNGR
- TARQEQ
- TARVOS
- TELESTO
- TETHYS
- **THRYM**
- THRYMR
- TITAN
- **UNK**
- YMIR

- **red** = misspelled.
- **violet** = needs correction.
- **yellow** = needs review and update.
- **orange** = not useful.
- Note: 25% of all Cassini ISS images are identified as TARGET_NAME = “SKY”.

On 10/21/19, 6:50 AM, "pds_operator@jpl.nasa.gov" <pds_operator@jpl.nasa.gov> wrote:

Name: Daniel Cordier
Email: daniel.cordier@univ-reims.fr
Type: Question
Comment: Dear All,
with colleagues we wrote a paper based on the interpretation of VIMS occultations data analyzed by Maltagliati et al. (2015) (particularly corresponding to T10, T78 and T53), the processed data were provided by Maltagliati (private com), but one of our Reviewer is asking why these data are not publicly available. I check in OPUS and aparently this is true, I wasn't able to find data corresponding to Maltagliati et al. (2015) (see their Table 1).
So, did I make a mistake? Any reason for the absence of this date in OPUS?
Many thanks in advance for any help.
Best regards,
Daniel Cordier

Location: https://tools.pds-rings.seti.org/opus/#/cols=opusid,instrument,planet,target,time1,observationduration

Dear Dr. Cordier,

It appears that these observations were identified as "TARGET=SUN", not "TARGET=TITAN", by the VIMS team. Here is a link to an OPUS query that returns all of the VIMS data files associated with Titan solar occultations:

https://tools.pds-rings.seti.org/opus/#/CASSINItargetcode=TI+(Titan)&instrument=Cassini+VIMS&target=Sun&cols=opusid,instrument,planet,target,time1,observationduration,CASSINIobsname&widgets=CASSINItargetcode,instrument,target&order=time1,opusid&view=search&browse=gallery&cart_browse=gallery&startobs=1&cart_startobs=1&detail=

I found them by specifying target=Sun but also using the "Cassini Target code", which the VIMS team used internally for their own purposes; in this case I have found that it does specify Titan.

I should add that we are working on improving our search capabilities for occultation data sets. As this case illustrates, the teams did not always provide the information one would need to find them easily.

Thank you for using OPUS!

Best, Mark Showalter

Needed: Reliable Metadata

- What’s do users expect from our metadata?
  - No surprises.
  - No need for “secret knowledge”.

Needed: Reliable Metadata

- What's do users expect from our metadata?
  - No surprises.
  - No need for "secret knowledge".

- Solution?
  - Can we demand more from our data providers? Probably not.
  - PDS Nodes will probably have to deal with these issues locally.
    - Do we need to coordinate our metadata updates across Discipline Nodes?

Needed: Reliable Metadata

- What’s do users expect from our metadata?
  - No surprises.
  - No need for “secret knowledge”.

- Solution?
  - Can we demand more from our data providers? Probably not.
  - PDS Nodes will probably have to deal with these issues locally.
    - Do we need to coordinate our metadata updates across Discipline Nodes?

- Note: If anything ever goes wrong, PDS will get the blame.

What do we need to do to support discovery?

(Finding data that you want but didn't know exists)

Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.

Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.
  - Example: Filter names

Voyager ISS

Filter Name [Voyager ISS]

Clear 19234 · UV 5681 · Violet 13093 · Blue 7394 · Green 13927 · Methane-U 772 · Sodium-D 144 · Orange 13426 · Methane-JST 2981

Cassini ISS

Filter [Cassini ISS]

BL1 21561 · BL1+GRN 438 · BL1+IR3 1 · BL1+MT3 1 · BL2 890 · BL2+P0 937 · BL2+P120 925 · BL2+P60 924 · CB1 3689 · CB1+HAL 1 · CB1+IRP0 3 · CB1+P0 922 · CB1+P120 921 · CB1+P60 922 · CB2 23270 · CB2+IRP0 2076 · CB2+IRP90 1417 · CB2+P0 785 · CB2+P120 818 · CB2+P60 820 · CB3 27802 · CB3+IRP0 1524 · CB3+IRP90 681 · CLEAR 207199 · GRN 19759 · GRN+IRP0 7 · GRN+P0 2126 · GRN+P120 2108 · GRN+P60 2112 · GRN+RED 516 · HAL 2136 · HAL+CB3 18 · IR1 9049 · IR1+CB2 1 · IR1+IR2 1763 · IR1+IRP0 44 · IR1+P0 4 · IR1+P120 4 · IR1+P60 4 · IR2 3470 · IR2+IR3 536 · IR2+IRP0 25 · IR2+IRP90 25 · IR3 6351 · IR3+IR4 355 · IR3+IRP0 165 · IR3+IRP90 29 · IR3+P0 1 · IR4 1631 · IR4+IRP0 116 · IR4+IRP90 117 · IR5 321 · IR5+IRP0 10 · IR5+IRP90 12 · IRP0 442 · IRP90 315 · MT1 9264 · MT1+IRP0 10 · MT1+P0 97 · MT1+P120 97 · MT1+P60 97 · MT2 16538 · MT2+IR1 2 · MT2+IRP0 1803 · MT2+IRP90 1172 · MT2+P0 627 · MT2+P120 619 · MT2+P60 618 · MT3 13486 · MT3+IRP0 927 · MT3+IRP90 743 · MT3+P120 16 · P0 241 · P120 153 · P60 227 · RED 16648 · RED+IR1 469 · RED+IR3 2 · RED+MT2 3 · UV1 3689 · UV1+BL2 2 · UV1+CB1 2 · UV1+CB2 2 · UV1+GRN 2 · UV1+IR1 2 · UV1+IR3 2 · UV1+UV3 2 · UV2 2256 · UV2+BL2 2 · UV2+CB1 2 · UV2+CB2 2 · UV2+GRN 2 · UV2+IR1 2 · UV2+IR3 2 · UV2+MT1 2 · UV2+UV3 136 · UV3 9082 · UV3+BL1 4 · UV3+HAL 2 · UV3+IR2 2 · UV3+IR4 2 · UV3+P0 1506 · UV3+P120 1485 · UV3+P60 1496 · UV3+RED 2 · VIO 7528 · VIO+CB3 1

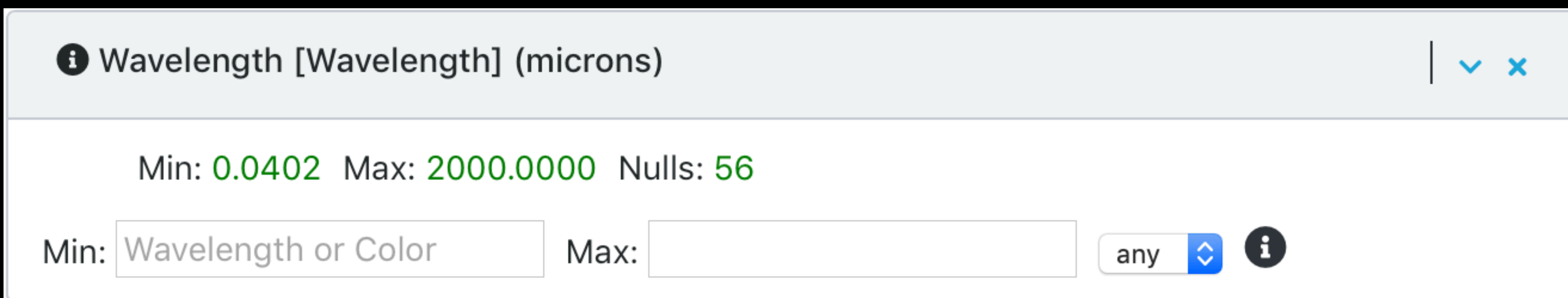
HST

Filter Name [Hubble]

ACS-CLEAR 348 · ACS-F115LP 2163 · ACS-F122M 3 · ACS-F125LP 2639 · ACS-F140LP 116 · ACS-F150LP 20 · ACS-F165LP 70 · ACS-F220W 55 · ACS-F220W+POL0UV 4 · ACS-F220W+POL120UV 4 · ACS-F220W+POL60UV 4 · ACS-F250W 74 · ACS-F250W+POL0UV 5 · ACS-F250W+POL120UV 5 · ACS-F250W+POL60UV 5 · ACS-F330W 83 · ACS-F330W+POL0UV 5 · ACS-F330W+POL120UV 5 · ACS-F330W+POL60UV 5 · ACS-F344N 27 · ACS-F435W 115 · ACS-F435W+POL0UV 5 · ACS-F435W+POL120UV 5 · ACS-F435W+POL60UV 5 · ACS-F475W 163 · ACS-F502N 46 · ACS-F550M 43 · ACS-F555W 66 · ACS-F606W 311 · ACS-F625W 7 · ACS-F625W+POL0V 3 · ACS-F625W+POL120V 3 · ACS-F625W+POL60V 3 · ACS-F658N 61 · ACS-F658N+POL0V 3 · ACS-F658N+POL120V 3 · ACS-F658N+POL60V 3 · ACS-F660N 1 · ACS-F775W 72 · ACS-F775W+POL0V 3 · ACS-F775W+POL120V 3 · ACS-F775W+POL60V 3 · ACS-F814W 152 · ACS-F850LP 29 · ACS-F892N 112 · ACS-FR601N 4 · ACS-FR656N 19 · ACS-FR782N 12 · ACS-FR914M 9 · ACS-FR931N 16 · ACS-PR130L 6 · NICMOS-BLANK 161 · NICMOS-F090M 18 · NICMOS-F095N 50 · NICMOS-F097N 49 · NICMOS-F108N 89 · NICMOS-F110M 59 · NICMOS-F110W 48 · NICMOS-F113N 24 · NICMOS-F145M 75 · NICMOS-F150W 4 · NICMOS-F160W 74 · NICMOS-F165M 132 · NICMOS-F166N 56 · NICMOS-F170M 7 · NICMOS-F171M 35 · NICMOS-F175W 4 · NICMOS-F180M 53 · NICMOS-F187N 130 · NICMOS-F187W 87 · NICMOS-F190N 75 · NICMOS-F204M 170 · NICMOS-F205W 19 · NICMOS-F207M 79 · NICMOS-F212N 131 · NICMOS-F215N 12 · NICMOS-F216N 6 · NICMOS-F222M 55 · NICMOS-F237M 173 · NICMOS-G096 13 · NICMOS-G141 13 · NICMOS-G206 143 · NICMOS-POL0L 12 · NICMOS-POL120L 12 · NICMOS-POL240L 12 · STIS-CLEAR 1374 · STIS-CRYSTAL QUARTZ 13 · STIS-LONG PASS 137 · STIS-LYMAN ALPHA 3 · STIS-ND3 2 · STIS-STRONTIUM FLUORIDE 285 · WFC3-F105W 2 · WFC3-F110W 8 · WFC3-F125W 3 · WFC3-F127M 16 · WFC3-F139M 92 · WFC3-F153M 16 · WFC3-F160W 31 · WFC3-F218W 8 · WFC3-F225W 57 · WFC3-F275W 144 · WFC3-F300X 6 · WFC3-F336W 13 · WFC3-F343N 31 · WFC3-F350LP 1405 · WFC3-F390M 4 · WFC3-F395N 57 · WFC3-F410M 14 · WFC3-F438W 259 · WFC3-F467M 197 · WFC3-F469N 2 · WFC3-F475W 8 · WFC3-F475X 6 · WFC3-F487N 2

Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.
  - Example: Filter names



Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.
    - Example: Filter names

Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.
  - Example: Filter names
- Consistent metadata is especially important for dataset-level searches.
  - PDS's repeated attempts at a single "grand solution" to categorize instruments and datasets have all failed.
  - PDS4's "facets" are a step in the right direction but are not sufficient.

Needed: Consistent Metadata

- Cross-mission/cross-instrument terminology is inconsistent and always will be.
    - Example: Filter names
- Consistent metadata is especially important for dataset-level searches.
    - PDS’s repeated attempts at a single “grand solution” to categorize instruments and datasets have all failed.
    - PDS4’s “facets” are a step in the right direction but are not sufficient.

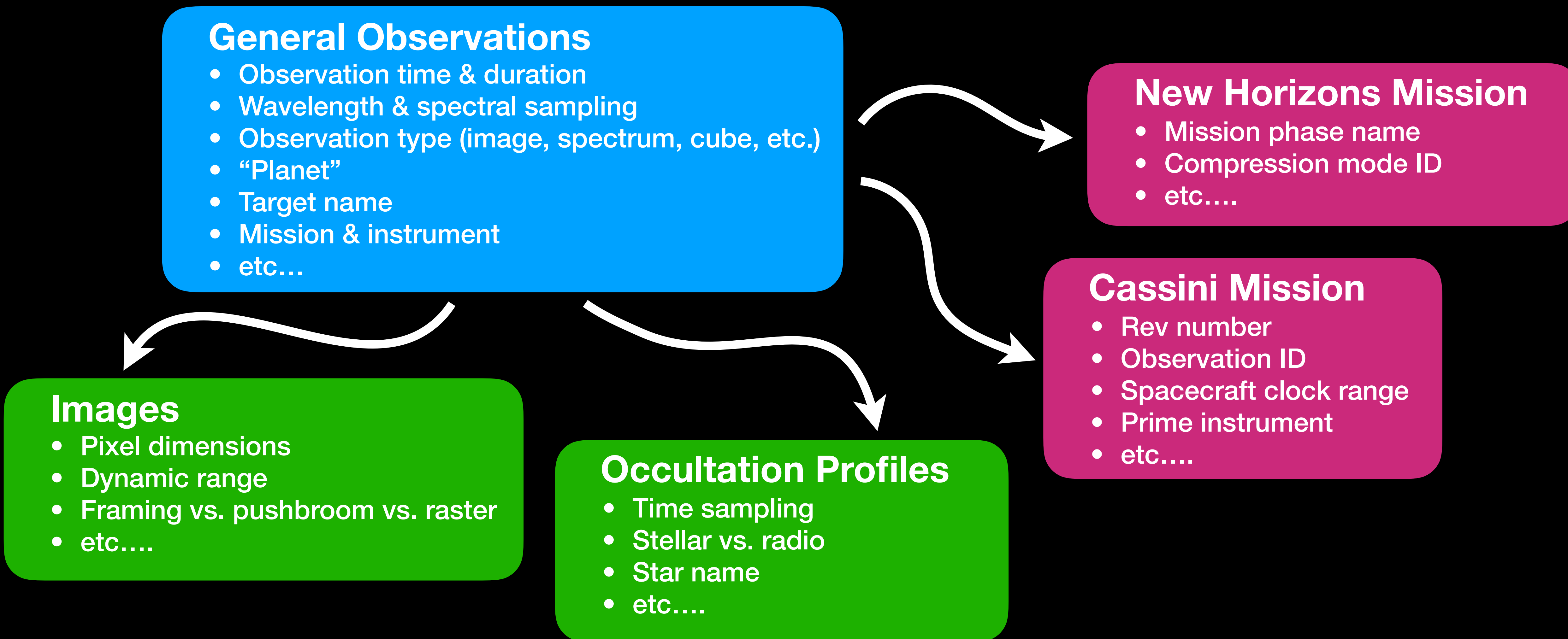
- Solution?
    - The OPUS import engine standardizes common fields across instruments and also defines new universal fields (e.g., wavelength).
    - OPUS uses a hierarchical, *object-oriented* schema with multiple inheritance to represent each product.

Object-Oriented Schema

General Observations

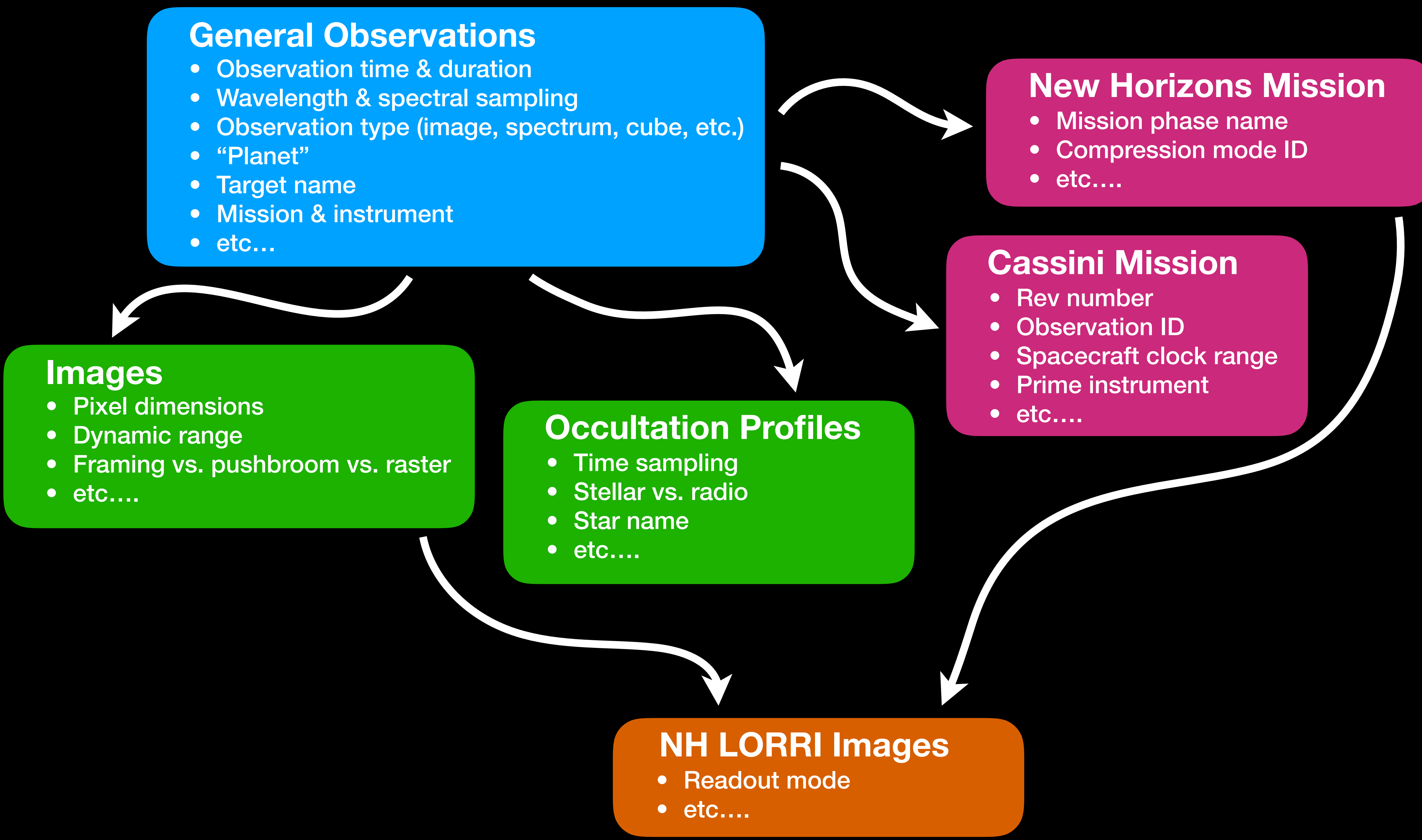
- Observation time & duration
- Wavelength & spectral sampling
- Observation type (image, spectrum, cube, etc.)
- "Planet"
- Target name
- Mission & instrument
- etc…

Object-Oriented Schema

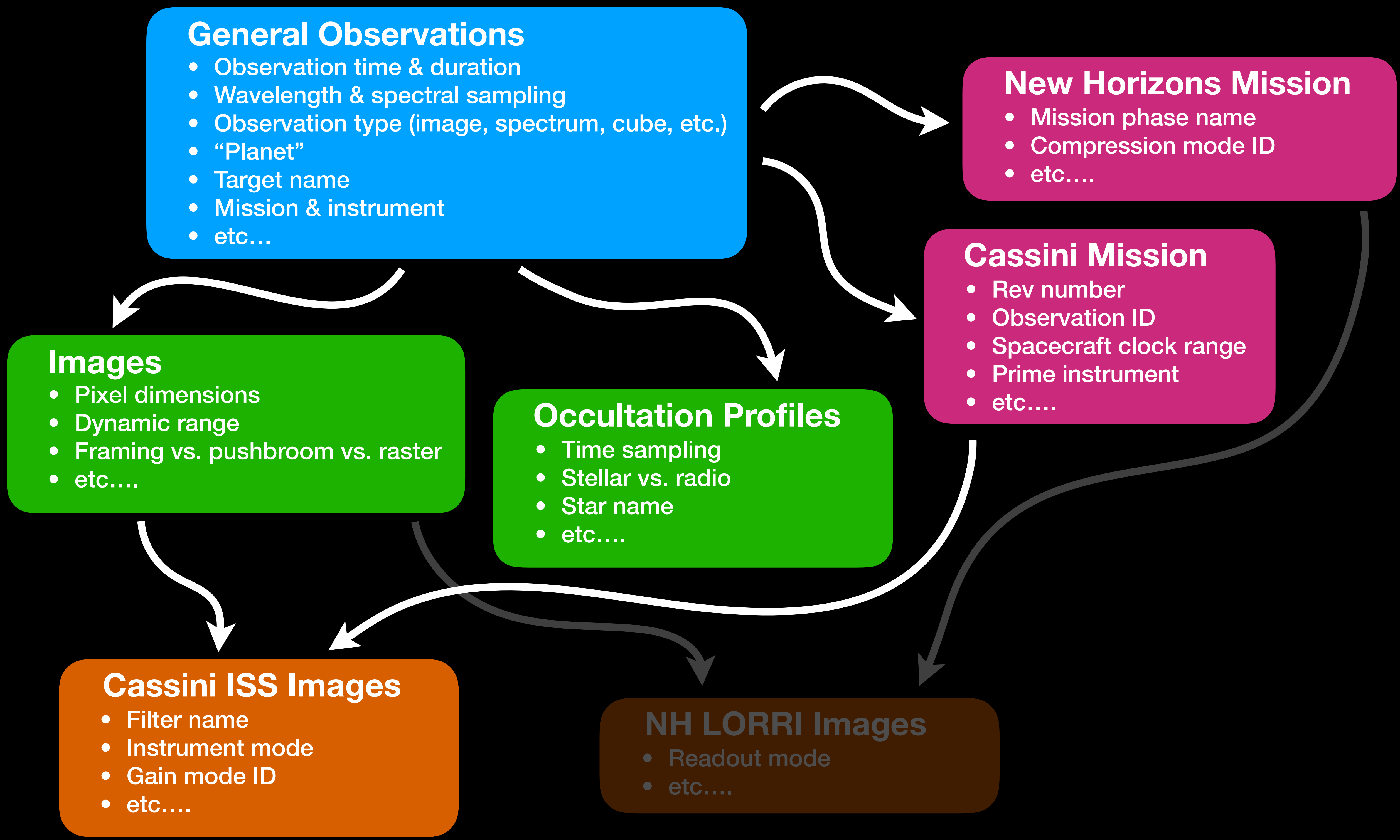


SETI INSTITUTE

Object-Oriented Schema



Object-Oriented Schema



Object-Oriented Schema

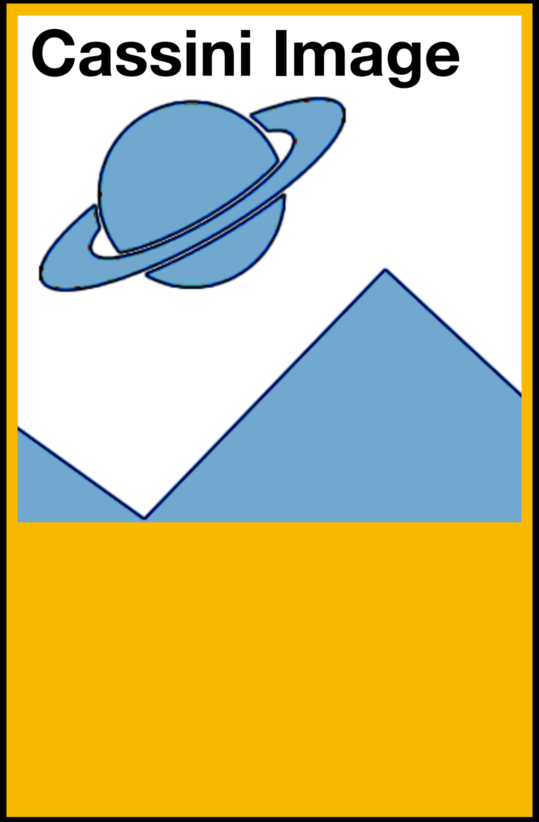
General Observations
- Observation time & duration
- Wavelength & spectral sampling
- **Observation type** (**image**, spectrum, cube, etc.)
- “Planet”
- Target name
- **Mission & instrument**
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- Pixel dimensions
- Dynamic range
- **Framing** vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….



Cassini ISS Images
- Filter name
- Instrument mode
- Gain mode ID
- etc….

Object-Oriented Schema

General Observations
- **Observation time** & duration
- Wavelength & spectral sampling
- Observation type (image, spectrum, cube, etc.)
- “Planet”
- Target name
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- Pixel dimensions
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini ISS Images
- Filter name
- Instrument mode
- Gain mode ID
- etc….

Cassini Image


START_TIME

Object-Oriented Schema

General Observations
- Observation time & duration
- **Wavelength & spectral sampling**
- Observation type (image, spectrum, cube, etc.)
- “Planet”
- Target name
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- Pixel dimensions
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini Image

START_TIME
FILTER_NAME

Cassini ISS Images
- **Filter name**
- Instrument mode
- Gain mode ID
- etc….

Object-Oriented Schema

General Observations
- Observation time & duration
- Wavelength & spectral sampling
- Observation type (image, spectrum, cube, etc.)
- "Planet"
- Target name
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- **Pixel dimensions**
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini Image

START_TIME
FILTER_NAME
INSTRUMENT_MODE

Cassini ISS Images
- Filter name
- **Instrument mode**
- Gain mode ID
- etc….

SETI INSTITUTE

Object-Oriented Schema

General Observations
- Observation time & duration
- Wavelength & spectral sampling
- Observation type (image, spectrum, cube, etc.)
- **“Planet”**
- **Target name**
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- **Rev number**
- **Observation ID**
- Spacecraft clock range
- **Prime instrument**
- etc….

Images
- **Pixel dimensions**
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini Image

START_TIME
FILTER_NAME
INSTRUMENT_MODE
OBSERVATION_ID

Cassini ISS Images
- Filter name
- **Instrument mode**
- Gain mode ID
- etc….

Search: Images of Jupiter

General Observations
- Observation time & duration
- Wavelength & spectral sampling
- **Observation type** (**image**, spectrum, cube, etc.)
- “Planet”
- **Target name**
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- Pixel dimensions
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini ISS Images
- Filter name
- Instrument mode
- Gain mode ID
- etc….

NH LORRI Images
- Readout mode
- etc….

SETI INSTITUTE

Search: Images of Jupiter in 2007

General Observations
- **Observation time** & duration
- Wavelength & spectral sampling
- **Observation type** (**image**, spectrum, cube, etc.)
- “Planet”
- **Target name**
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc.…

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc.…

Images
- Pixel dimensions
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc.…

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc.…

Cassini ISS Images
- Filter name
- Instrument mode
- Gain mode ID
- etc.…

NH LORRI Images
- Readout mode
- etc.…

Search: Images of Jupiter in 2007

General Observations
- **Observation time** & duration
- Wavelength & spectral sampling
- **Observation type** (**image**, spectrum, cube, etc.)
- "Planet"
- **Target name**
- Mission & instrument
- etc…

New Horizons Mission
- Mission phase name
- Compression mode ID
- etc….

Cassini Mission
- Rev number
- Observation ID
- Spacecraft clock range
- Prime instrument
- etc….

Images
- Pixel dimensions
- Dynamic range
- Framing vs. pushbroom vs. raster
- etc….

Occultation Profiles
- Time sampling
- Stellar vs. radio
- Star name
- etc….

Cassini ISS Images
- Filter name
- Instrument mode
- Gain mode ID
- etc….

NH LORRI Images
- Readout mode
- etc….

Object-Oriented Schema

- This approach provides the best of both worlds:
  - High-level search is available across all data sets.
  - All instrument- and mission-specific parameters are available once a search has been narrowed down to a single dataset.
  - The transition between dataset-level search and product-level search is seamless.
- The “flat” tables delivered to PDS require serious modification in order to support this approach.
  - …raising the question: Will the PDS4 registry really be able to support data discovery the way we would like?

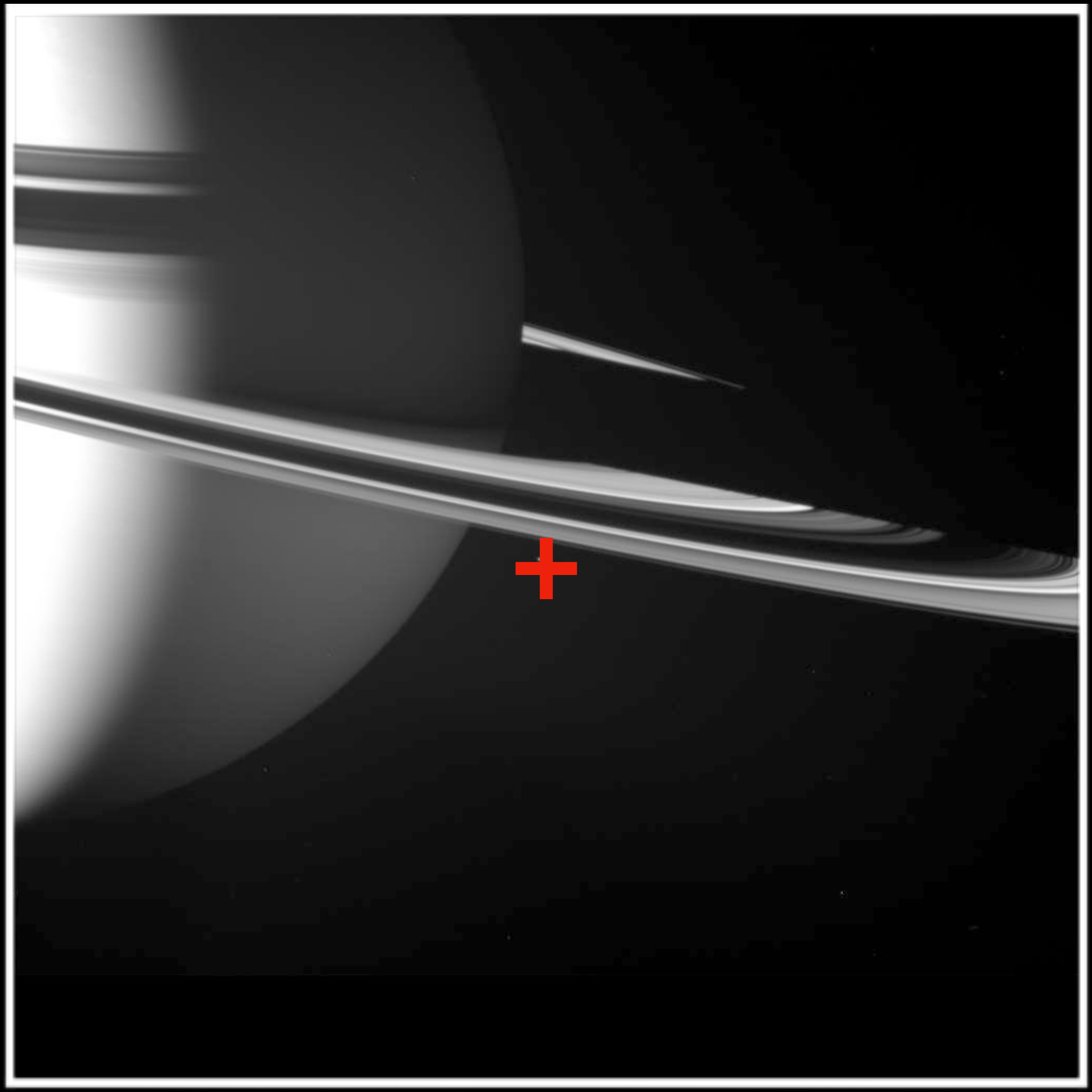
What do we need to do to support exploration?

(Finding data that exists but you didn't know you want)

Needed: Detailed Metadata

The problem:

- Non-mapping instrument teams rarely provide metadata suitable for robust geometric searches.
- Information may be provided for the center of the field of view, or perhaps for the four corners.
- Quantities like sub-spacecraft latitude and longitude do not tell you where the instrument was pointed.
- Robust geometric search requires a comprehensive sampling of the field of view.



Needed: Detailed Metadata

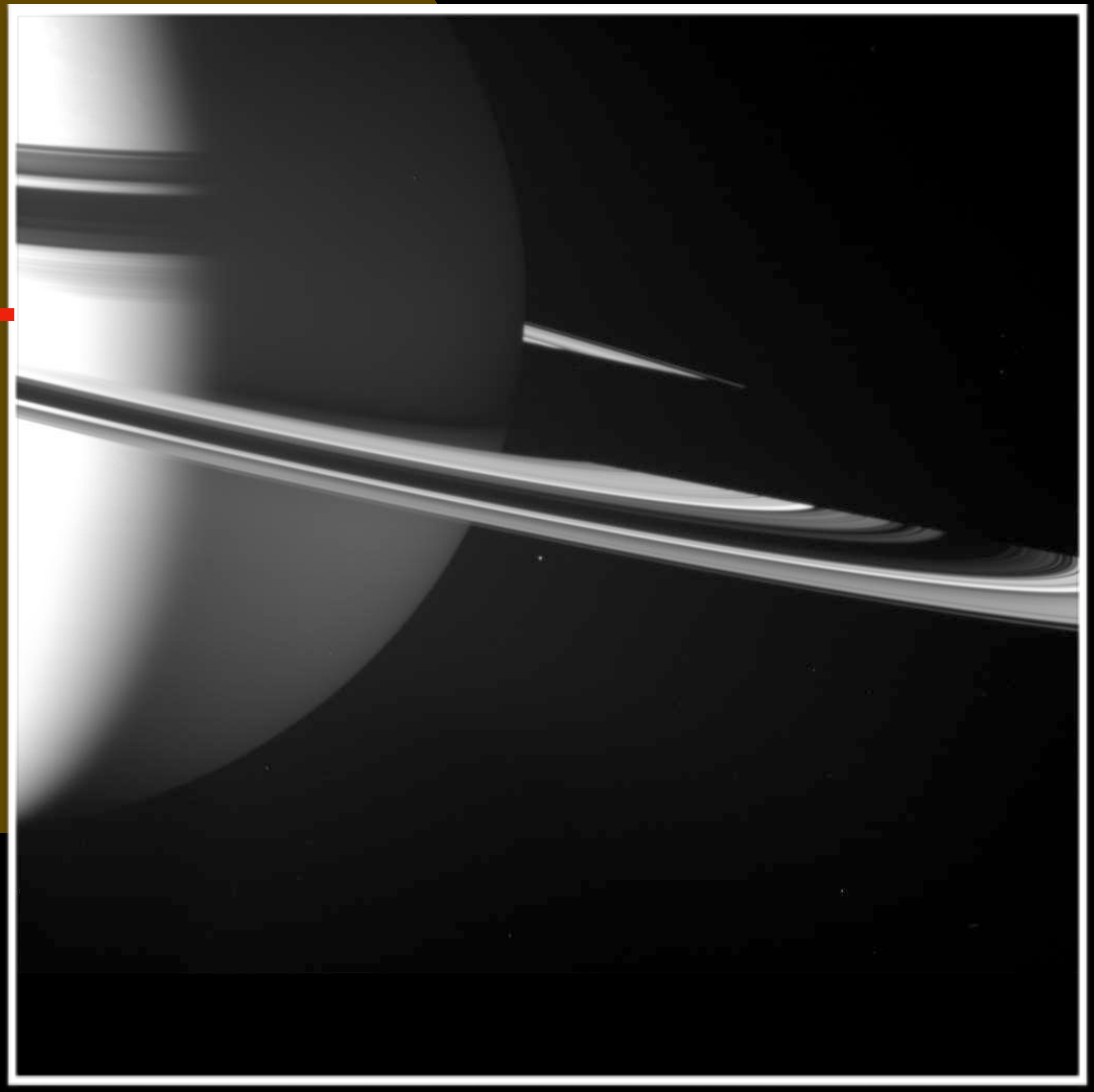
The problem:

- Non-mapping instrument teams rarely provide metadata suitable for robust geometric searches.
- Information may be provided for the center of the field of view, or perhaps for the four corners.
- Quantities like sub-spacecraft latitude and longitude do not tell you where the instrument was pointed.
- Robust geometric search requires a comprehensive sampling of the field of view.

Needed: Detailed Metadata

The problem:

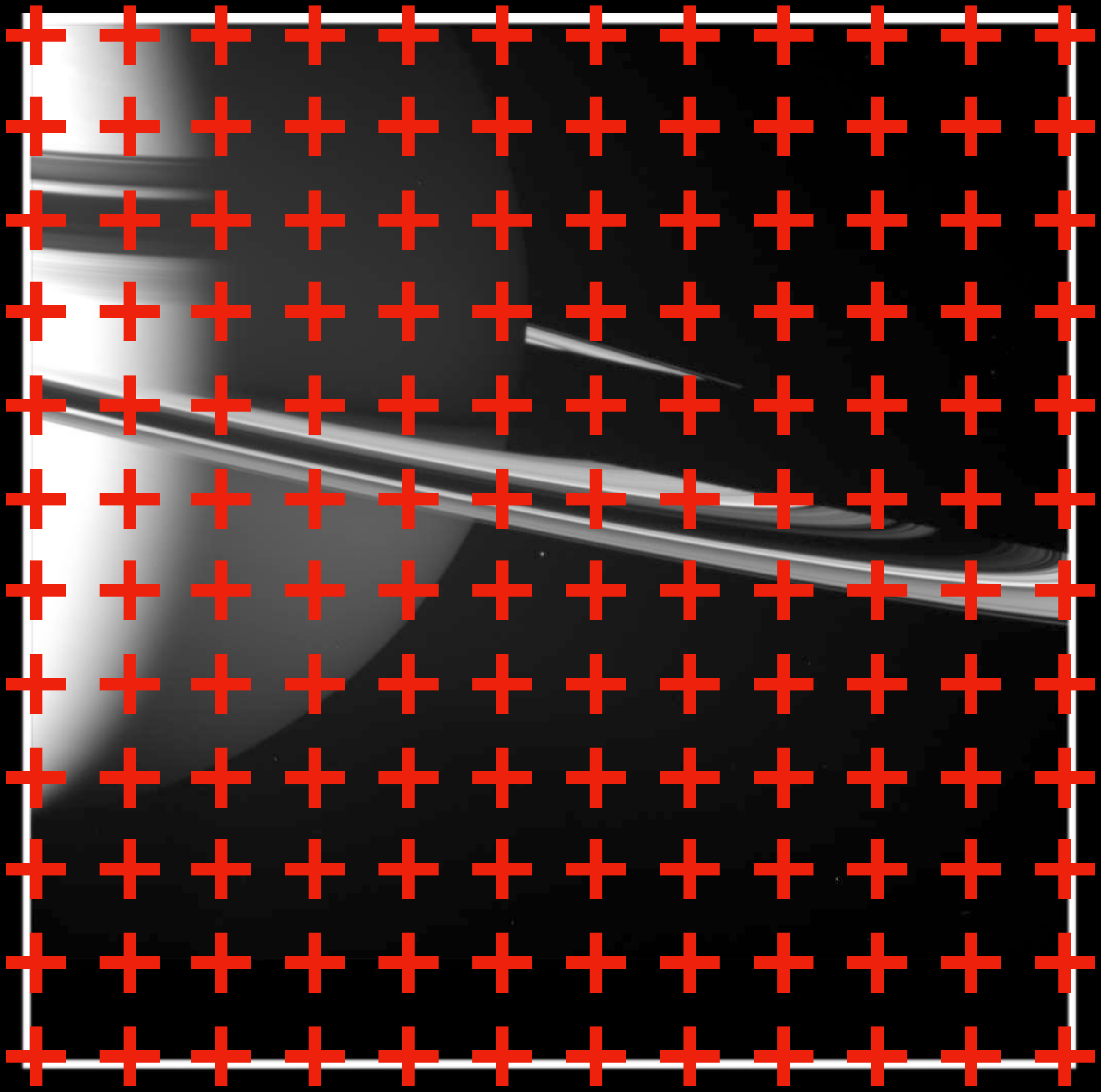
- Non-mapping instrument teams rarely provide metadata suitable for robust geometric searches.
- Information may be provided for the center of the field of view, or perhaps for the four corners.
- Quantities like sub-spacecraft latitude and longitude do not tell you where the instrument was pointed.
- Robust geometric search requires a comprehensive sampling of the field of view.



Needed: Detailed Metadata

The problem:

- Non-mapping instrument teams rarely provide metadata suitable for robust geometric searches.
- Information may be provided for the center of the field of view, or perhaps for the four corners.
- Quantities like sub-spacecraft latitude and longitude do not tell you where the instrument was pointed.
- Robust geometric search requires a comprehensive sampling of the field of view.



Needed: Sensible Granularity

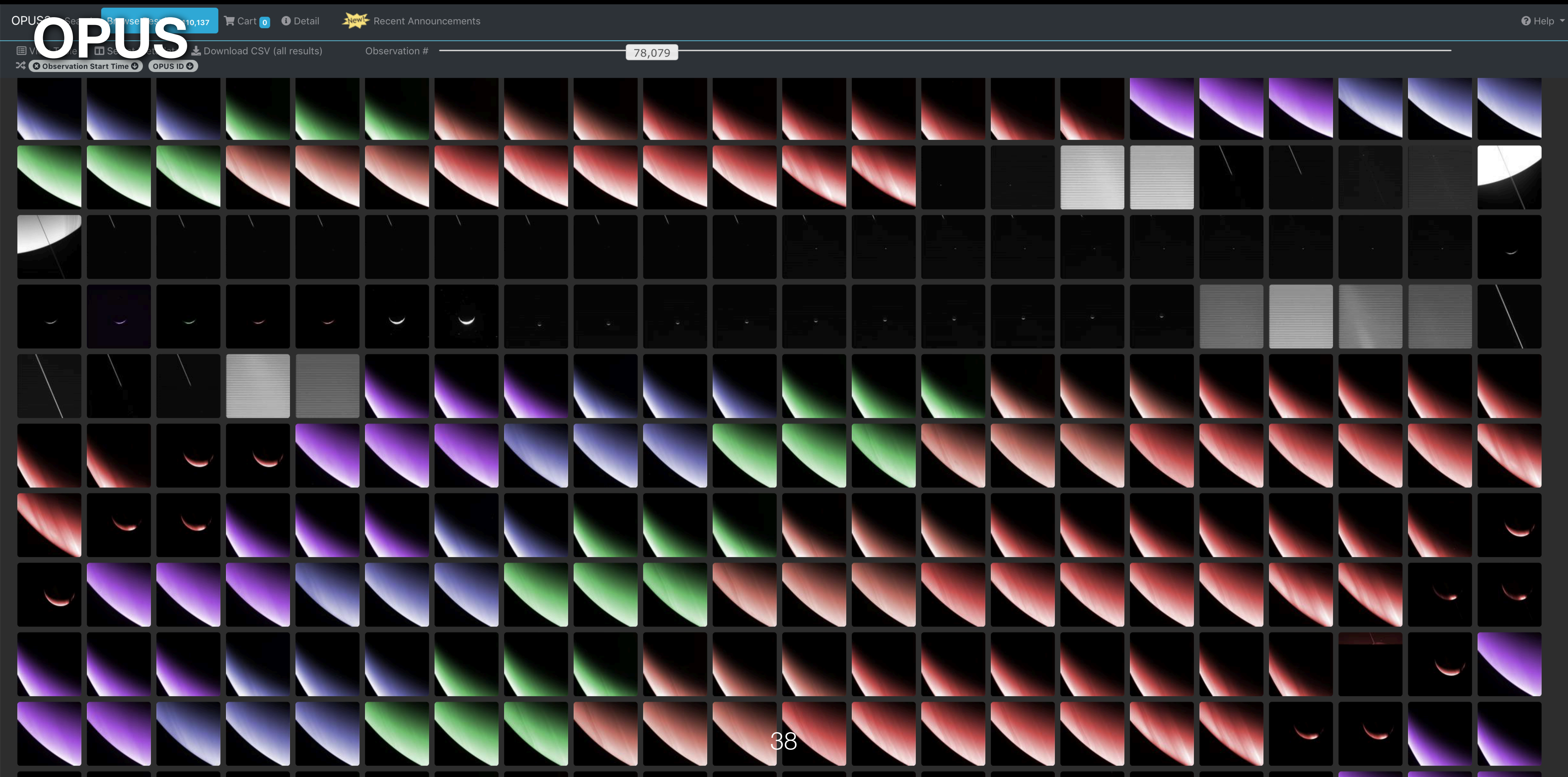
- PDS4 organization:
  - Bundle (e.g., all Cassini images) ⤵
    - Collection (e.g., raw or calibrated or browse) ⤵
      - Product (e.g., image1 or image2 or image3)

- Better for discovery and exploration:
  - Bundle (e.g., all Cassini images) ⤵
    - "Observation" (e.g., image1 or image2 or image3) ⤵
      - Product (e.g., raw or calibrated or browse)

- "Observation" = a selectable unit of data.
- In OPUS, users search for observations. Only after a query is complete do they decide what kinds of products to download.
- The PDS organization around products rather than observations clouds our thinking and interferes with the user experience.
  - The PDS4 mechanism for defining associations between products helps but is not sufficient.

Needed: Informative Browse Products

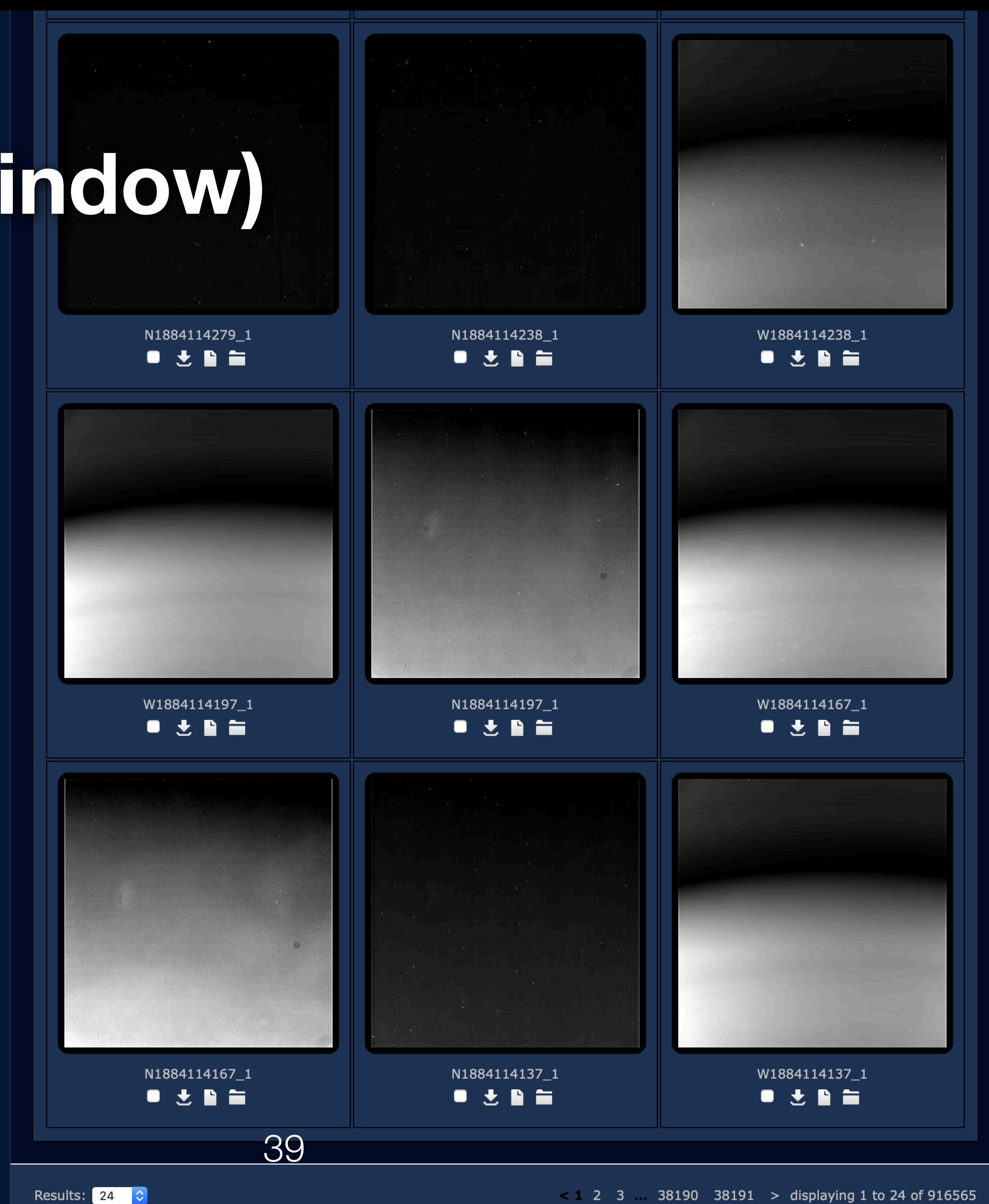
- Well-crafted browse products:
    - should make it possible to review thousands of data products quickly.



Needed: Informative Browse Products

- Well-crafted browse products:
  - should make it possible to review thousands of data products quickly.

Image Atlas
(same size browser window)



Needed: <u>Informative</u> Browse Products

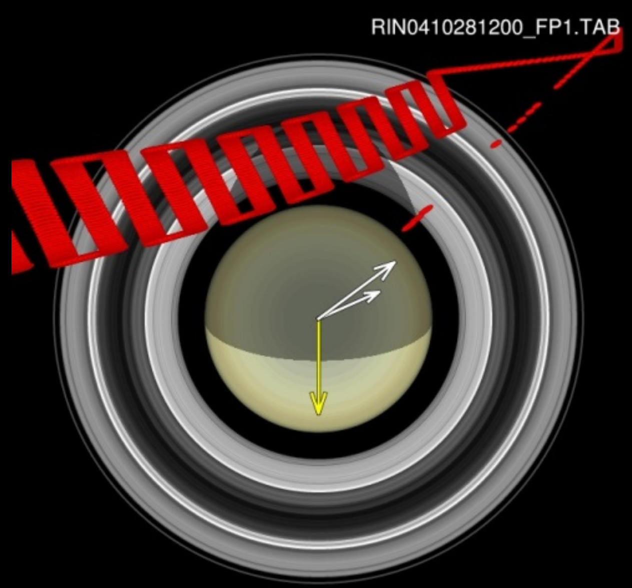
- Well-crafted browse products:
  - should make it possible to review thousands of data products quickly.

Needed: <u>Informative</u> Browse Products

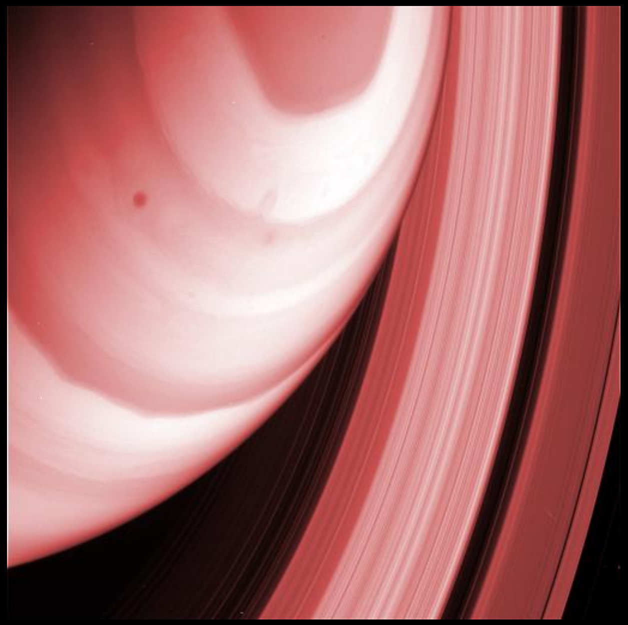
- Well-crafted browse products:
  - should make it possible to review thousands of data products quickly.
  - are critical to supporting discovery and (especially) exploration.
  - are not easy!

Needed: Informative Browse Products

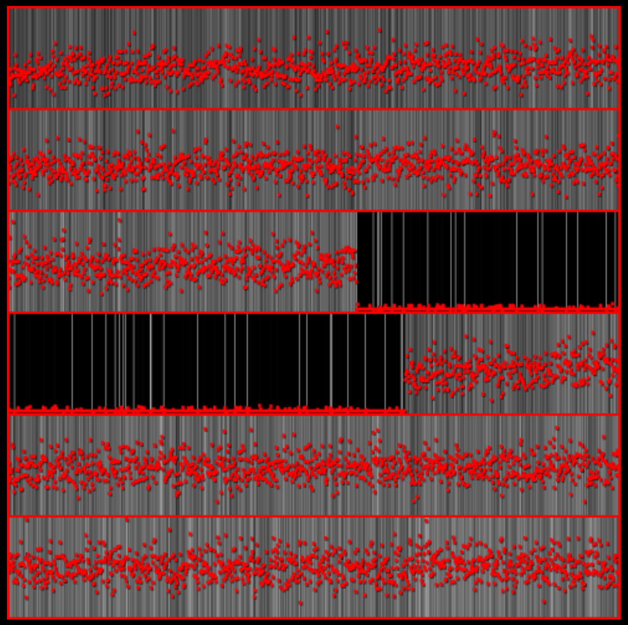
- Well-crafted browse products:
  - should make it possible to review thousands of data products quickly.
  - are critical to supporting discovery and (especially) exploration.
  - are not easy!



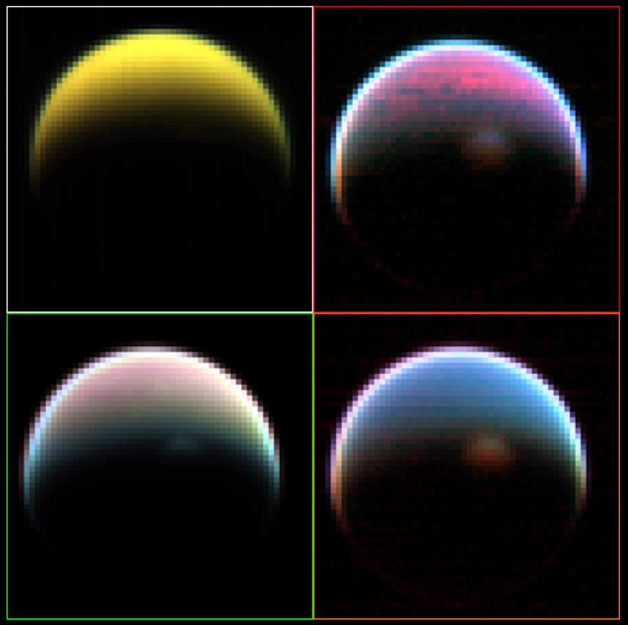
Cassini CIRS footprints



Cassini ISS, RED filter



Cassini UVIS time series

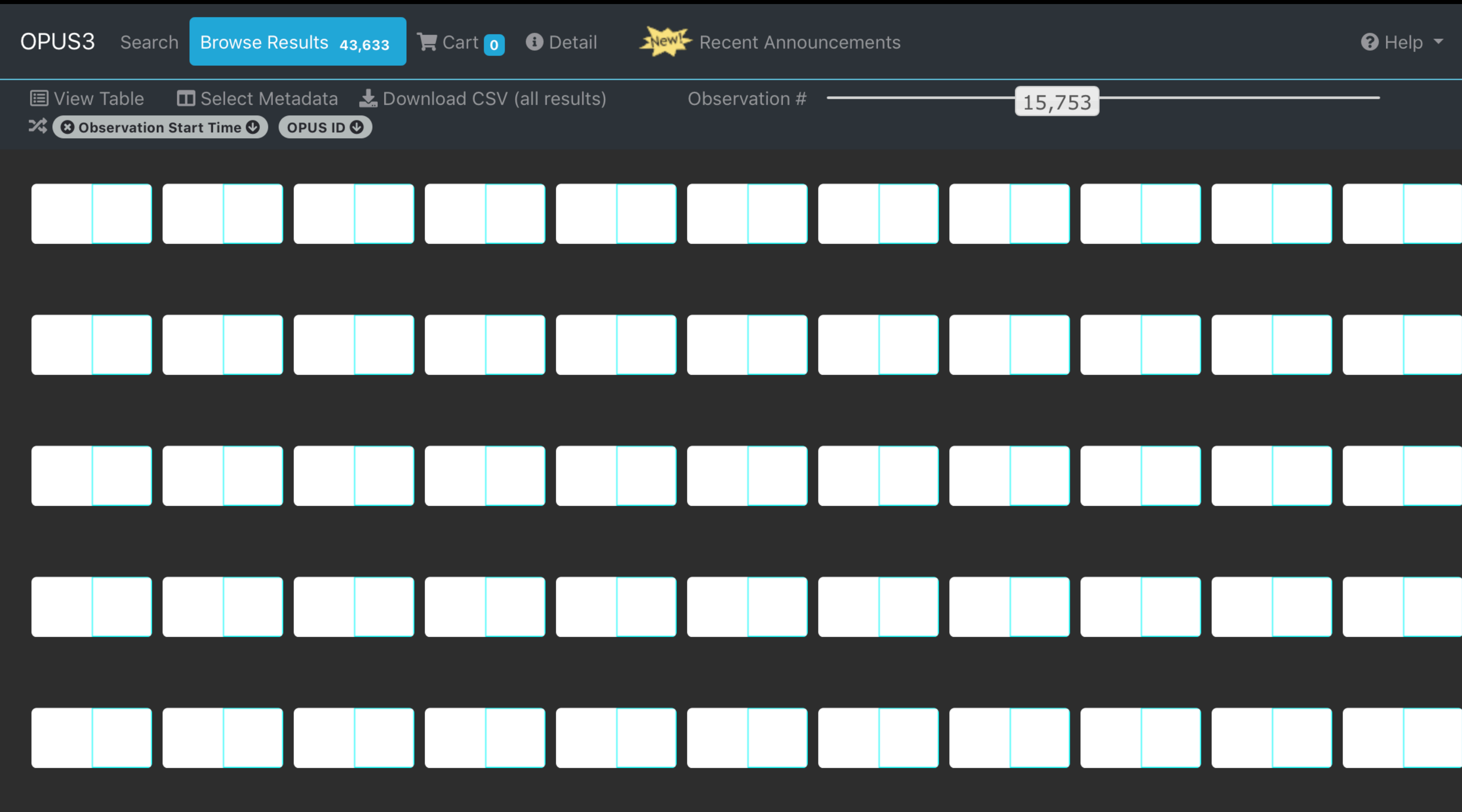


Cassini VIMS Visual/IR, Titan clouds

SETI INSTITUTE

Needed: Informative Browse Products

Dirty Laundry: Our previews of VIMS solar occultation data products… still need some work



Ten Years of OPUS: Lessons Learned

- OPUS supports Search, Discovery and Exploration across key outer planets data sets.
- Team-delivered metadata can be very unreliable and difficult to repair.
- To enable cross-mission, cross-instrument search capabilities, we need to:
  - identify inevitable inconsistencies in terminology and define “universal” quantities to supplement them.
  - implement our databases as hierarchical, “object-oriented” schemas rather than as “flat” tables.
- We probably cannot expect missions and instrument teams to provide the detailed, geometric metadata that are needed for serious exploration.
  - Discipline Nodes have an important role to play here.
- The most sensible “searchable unit” of data is not necessarily the product.
- High-quality, informative browse products are a critical and under-appreciated component of data discovery and exploration.
- Coming next: Data exploration can be and should be fun.