

Enabling Large-Scale Dataset Analytics: Hierarchical Adaptive Tiling Scheme (HATS)

Mario Juric

DiRAC Institute Director | LINCC FW
Professor of Astronomy, University of Washington

with Wilson Bebe, Doug Branton, Sandro Campos, Neven Caplar, Melissa DeLucchi, Derek Jones, Jeremy Kubica, Kostya Malanchev, Sean McGuire, Colin Slater, Steven Stetzler, Max West, Sam Wyatt



DATA INTENSIVE RESEARCH IN
ASTROPHYSICS AND COSMOLOGY



LINCC



AST-2003196

and the LINCC Frameworks Analytics Group



The LINCC Frameworks Project

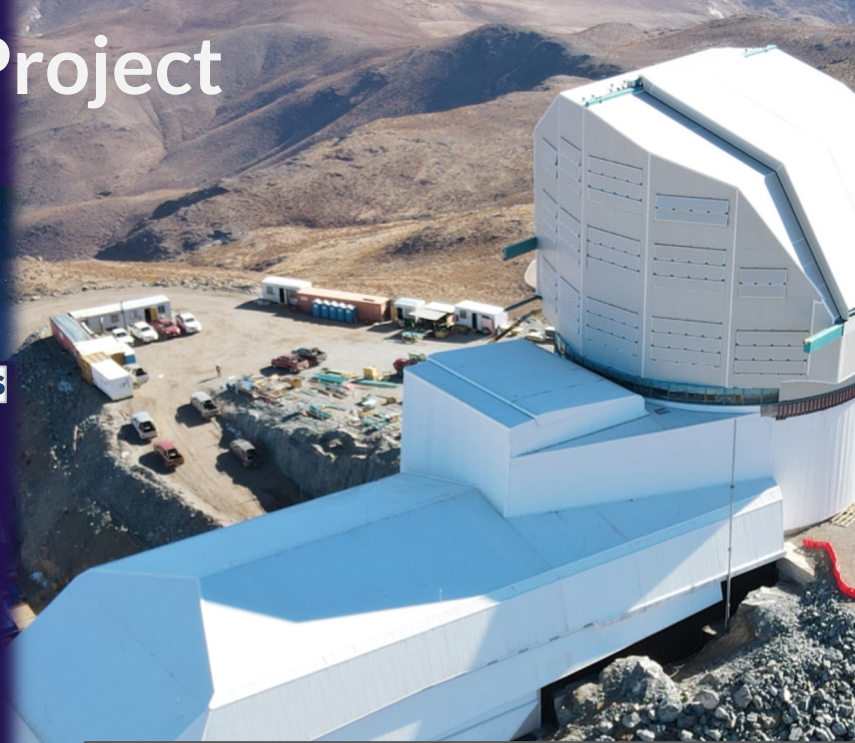
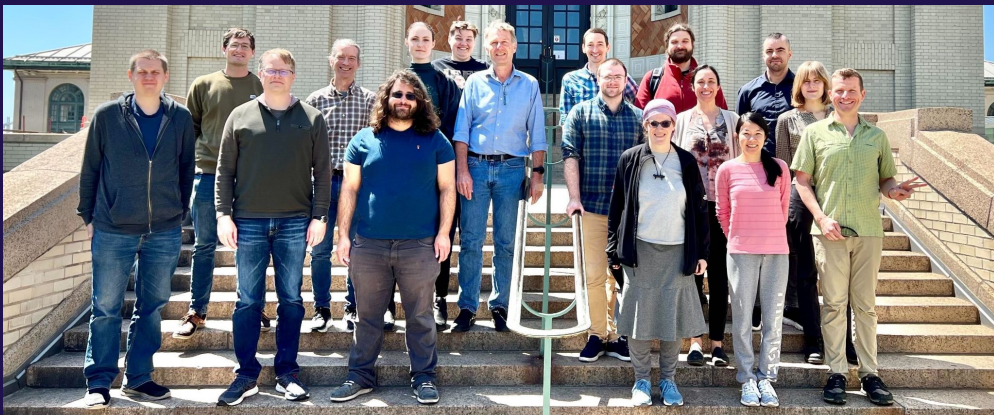
LSST Interdisciplinary Network For Collaboration And Computing

To collaboratively develop open computing systems and algorithms needed for large-survey analyses.



Two LINCC-FW hubs:

- Carnegie Mellon University
- University of Washington



The Legacy Survey of Space and Time

Deep synoptic optical survey, coming in 2025.

Repeated imaging of the visible sky to ~24th mag.

10 years of operation.

60 PB of raw data.

40 billion stars, galaxies, asteroids.

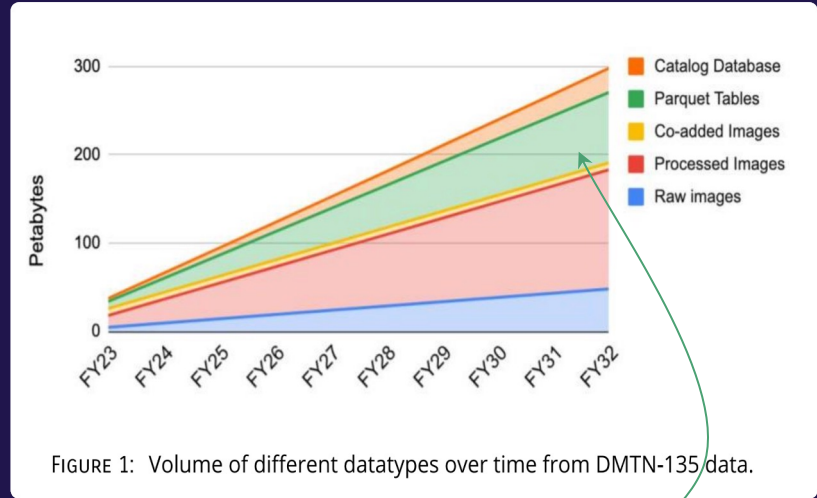
30 trillion observations.

Objective

Develop a system (format, tools) to deliver and enable end-user analyses on 10TB+ (catalog) datasets

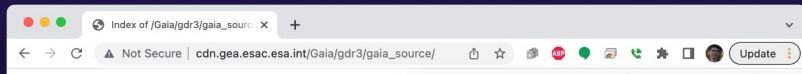
Large-dataset Analytics: Partitioned Files

- Relational databases are not ideal for this type of work. Poor UX, too many bottlenecks.
- Industry state-of-the-art is to use distributed analytics tools operating on files.
- Distributed computation achieved through partitioning.

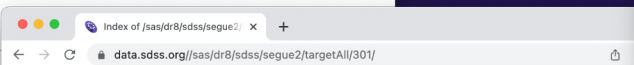


```
import pandas as pd
pd.read_parquet('example_pa.parquet', engine='pyarrow')
```

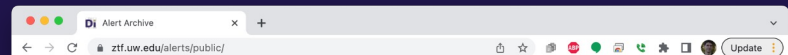
How to partition? Historically, we haven't generally given this much thought...



File Name	File Size
GaiaSource_000000-003111.csv.gz	05-May-2022
GaiaSource_003112-005263.csv.gz	05-May-2022
GaiaSource_005264-006601.csv.gz	05-May-2022
GaiaSource_006602-007952.csv.gz	05-May-2022
GaiaSource_007953-010234.csv.gz	05-May-2022
GaiaSource_010235-012597.csv.gz	05-May-2022
GaiaSource_012598-014045.csv.gz	05-May-2022
GaiaSource_014046-015169.csv.gz	05-May-2022
GaiaSource_015170-016240.csv.gz	05-May-2022
GaiaSource_016241-017018.csv.gz	05-May-2022
GaiaSource_017019-017658.csv.gz	05-May-2022
GaiaSource_017659-018028.csv.gz	05-May-2022
GaiaSource_018029-018477.csv.gz	05-May-2022
GaiaSource_018478-019161.csv.gz	05-May-2022
GaiaSource_019162-019657.csv.gz	05-May-2022
GaiaSource_019658-020091.csv.gz	05-May-2022
GaiaSource_020092-020493.csv.gz	05-May-2022
GaiaSource_020494-020767.csv.gz	05-May-2022
GaiaSource_020768-020984.csv.gz	05-May-2022
GaiaSource_020985-021233.csv.gz	05-May-2022
GaiaSource_021234-021411.csv.gz	05-May-2022
GaiaSource_021412-021665.csv.gz	05-May-2022
GaiaSource_021666-021919.csv.gz	05-May-2022
GaiaSource_021920-022158.csv.gz	05-May-2022
GaiaSource_022159-022410.csv.gz	05-May-2022
GaiaSource_022411-022698.csv.gz	05-May-2022
GaiaSource_022699-022881.csv.gz	05-May-2022
GaiaSource_022882-023058.csv.gz	05-May-2022
GaiaSource_023059-023264.csv.gz	05-May-2022
GaiaSource_023265-023450.csv.gz	05-May-2022
GaiaSource_023451-023649.csv.gz	05-May-2022
GaiaSource_023650-023910.csv.gz	05-May-2022
GaiaSource_023911-024205.csv.gz	05-May-2022
GaiaSource_024206-024526.csv.gz	05-May-2022
GaiaSource_024527-025166.csv.gz	05-May-2022
GaiaSource_025167-025691.csv.gz	05-May-2022
GaiaSource_025692-026057.csv.gz	05-May-2022
GaiaSource_026058-026390.csv.gz	05-May-2022
GaiaSource_026391-026648.csv.gz	05-May-2022
GaiaSource_026649-027106.csv.gz	05-May-2022
GaiaSource_027107-027517.csv.gz	05-May-2022
GaiaSource_027518-027832.csv.gz	05-May-2022
GaiaSource_027833-028076.csv.gz	05-May-2022
GaiaSource_028077-028318.csv.gz	05-May-2022



File Name	File Size
Parent directory/	-
1000/	-
1006/	-
1009/	-
1010/	-
1011/	-
1013/	-
1022/	-
1024/	-
1033/	-
1035/	-
1037/	-
1040/	-
1043/	-
1045/	-
1055/	-
1056/	-
1057/	-
109/	-
1119/	-
1120/	-
1122/	-
1133/	-
1140/	-
1142/	-
1231/	-
1233/	-
1239/	-
1241/	-
125/	-
1302/	-
1329/	-
1331/	-



What is included?

Below you will find compressed tar archives of ZTF event alerts (observations detected in image differences). Each tar file contains alerts collected in the given night (UTC-based), with each alert stored in a separate file in the AVRO format. To get you started, we offer a repository with low basic utilities for reading AVRO, serialized data, as well as an example Jupyter notebook. The schema fields are described here.

Why this service?

We are providing this archive as simple alternative to public event brokers. Full-featured event brokers that provide real-time access to these alerts include MARS, Lasair, ANTARES, and ALERCE.

Known caveats

- The data provided on this site is generated automatically. The files provided contain a full, unfiltered, 5-sigma alert stream. Depending on your science case, you may wish to improve the purity of your sample by filtering the data on the included attributes such as the signal-to-noise ratio or the real-bogus score.
- Users interested in un-subtracted archival photometry should consider the ZTF Data Releases, accessible at IRSA.
- A subset of events obtained through Caltech time are made public here in "programid3" tarballs; as of this writing these are additional observations of the current TESS sector.

Name	Last modified	Size
ztf_public_20220810.tar.gz	10 hours ago	8.3G
ztf_public_20220809.tar.gz	1 day ago	8.3G
ztf_public_20220808.tar.gz	2 days ago	1.7G
ztf_public_20220807.tar.gz	3 days ago	10G
ztf_public_20220806.tar.gz	4 days ago	11G
ztf_public_20220805.tar.gz	5 days ago	3.6G
ztf_public_20220804.tar.gz	5 days ago	6.3G
ztf_public_20220803.tar.gz	7 days ago	11G
ztf_public_20220802.tar.gz	8 days ago	17G
ztf_public_20220801.tar.gz	9 days ago	3.4G
ztf_public_20220731.tar.gz	10 days ago	2.8G
ztf_public_20220730.tar.gz	11 days ago	14G

2012-Dec-19 10:33	2012-Dec-19 10:33	2012-Dec-19 10:33
2012-Dec-19 10:33	2012-Dec-19 10:33	2012-Dec-19 10:33
2012-Dec-19 10:33	2012-Dec-19 10:33	2012-Dec-19 10:33
2012-Dec-19 10:34	2012-Dec-19 10:34	2012-Dec-19 10:34
2012-Dec-19 10:34	2012-Dec-19 10:34	2012-Dec-19 10:34
2012-Dec-19 10:34	2012-Dec-19 10:34	2012-Dec-19 10:34

There's value in thinking this through, and standardizing

- Users know what to expect and how to handle the dataset
- High quality, shared, analytics tools can be written
- Pre-staging/ETL may be avoided
- Multi-dataset analytics can supported

- Bulk export files == bulk analytics files
- Easier to generate and support for providers
- Can share code and infrastructure (e.g. mirroring, caching)

Think of all the wonderful tools and ecosystems that sprung up around HiPS!

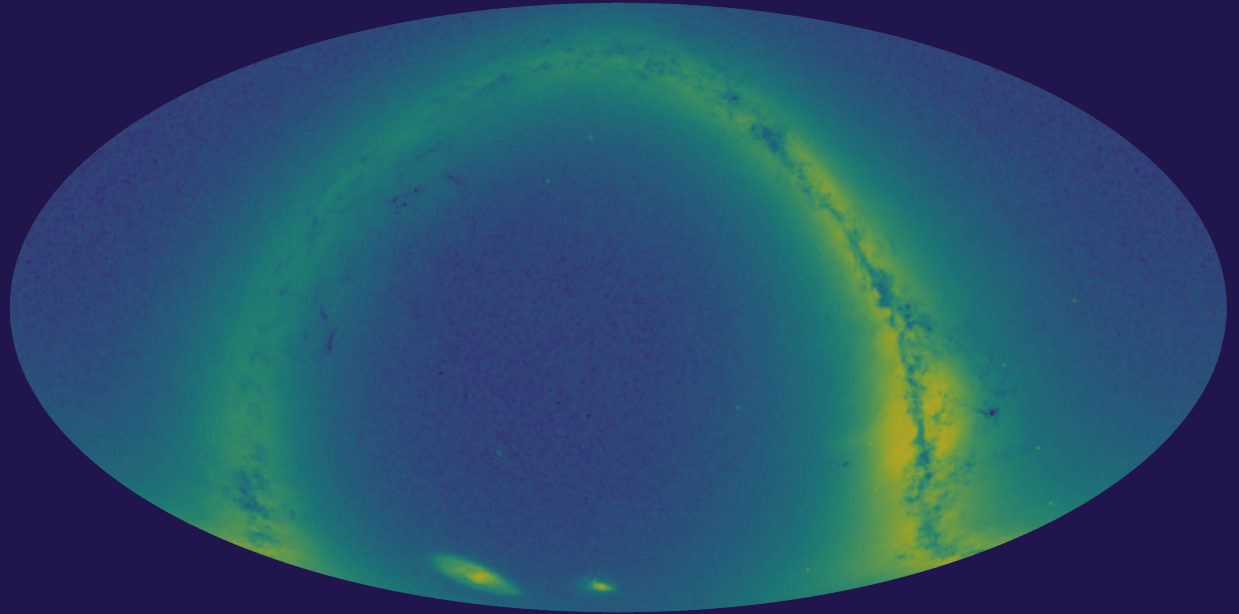


Introducing HATS: Hierarchical Adaptive Tiling Scheme

Gaia DR3

The 1.8Bn sources
released in Gaia DR3

A single ASCII file would
be about ~680GB in size,
(gzip compressed).



1. Partitioning: HEALPix

Partition the sky into NSIDE=1 (order=0)
HEALPix tiles, map tiles to files.

Example:

```
Norder0-Npix0.csv.gz  
Norder0-Npix1.csv.gz  
Norder0-Npix2.csv.gz  
Norder0-Npix3.csv.gz  
Norder0-Npix4.csv.gz  
Norder0-Npix5.csv.gz  
Norder0-Npix6.csv.gz  
Norder0-Npix7.csv.gz  
Norder0-Npix8.csv.gz  
Norder0-Npix9.csv.gz  
Norder0-Npix10.csv.gz  
Norder0-Npix11.csv.gz
```



Problem: Severely unbalanced file sizes

Pixel 4 (Galactic pole) ~ 20GB

Pixel 10 (Galactic center) ~ 400GB.

Simple file-based parallelization fails.

Example

Norder0-Npix0.csv.gz

Norder0-Npix1.csv.gz

Norder0-Npix2.csv.gz

Norder0-Npix3.csv.gz

Norder0-Npix4.csv.gz

Norder0-Npix5.csv.gz

Norder0-Npix6.csv.gz

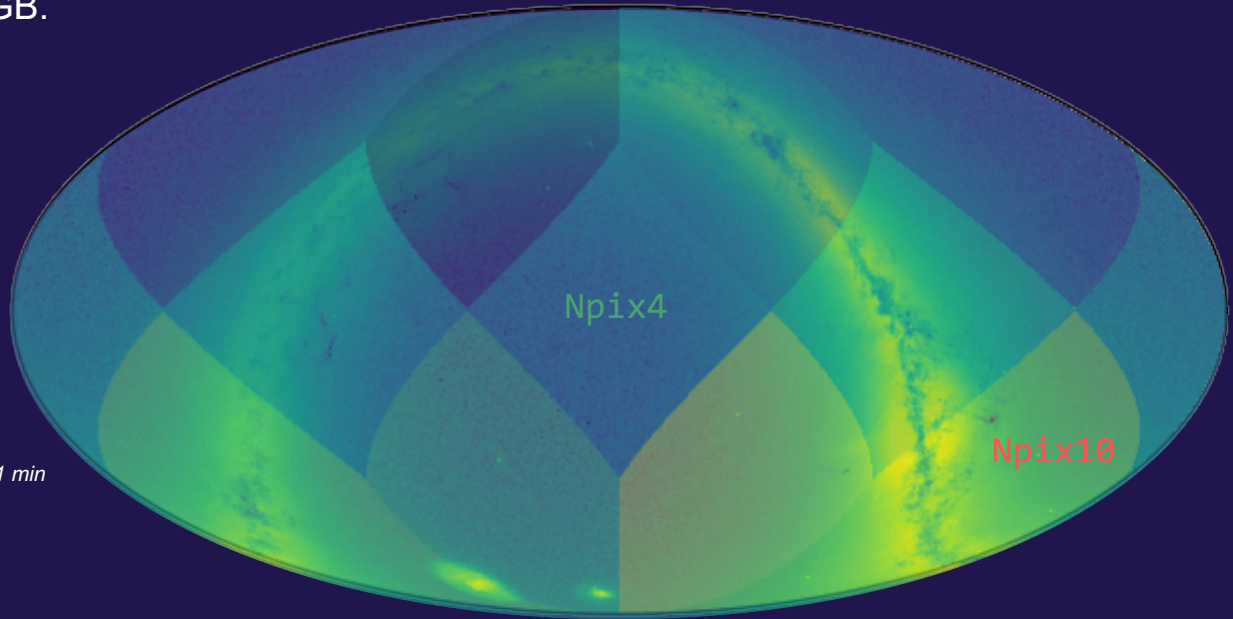
Norder0-Npix7.csv.gz

Norder0-Npix8.csv.gz

Norder0-Npix9.csv.gz

Norder0-Npix10.csv.gz

Norder0-Npix11.csv.gz

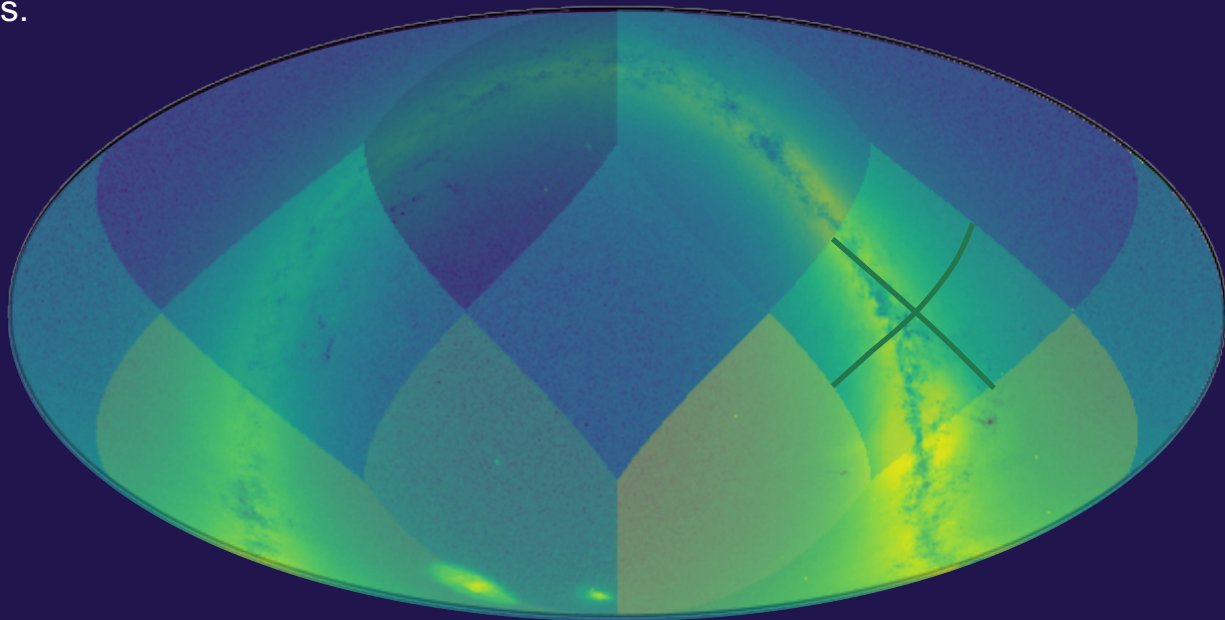


Solution: Partition Hierarchically

If too many sources fall into a pixel,
split it into four higher order pixels.

Example

Norder0-Npix0.csv.gz
...
Norder0-Npix7.csv.gz
...
Norder0-Npix11.csv.gz

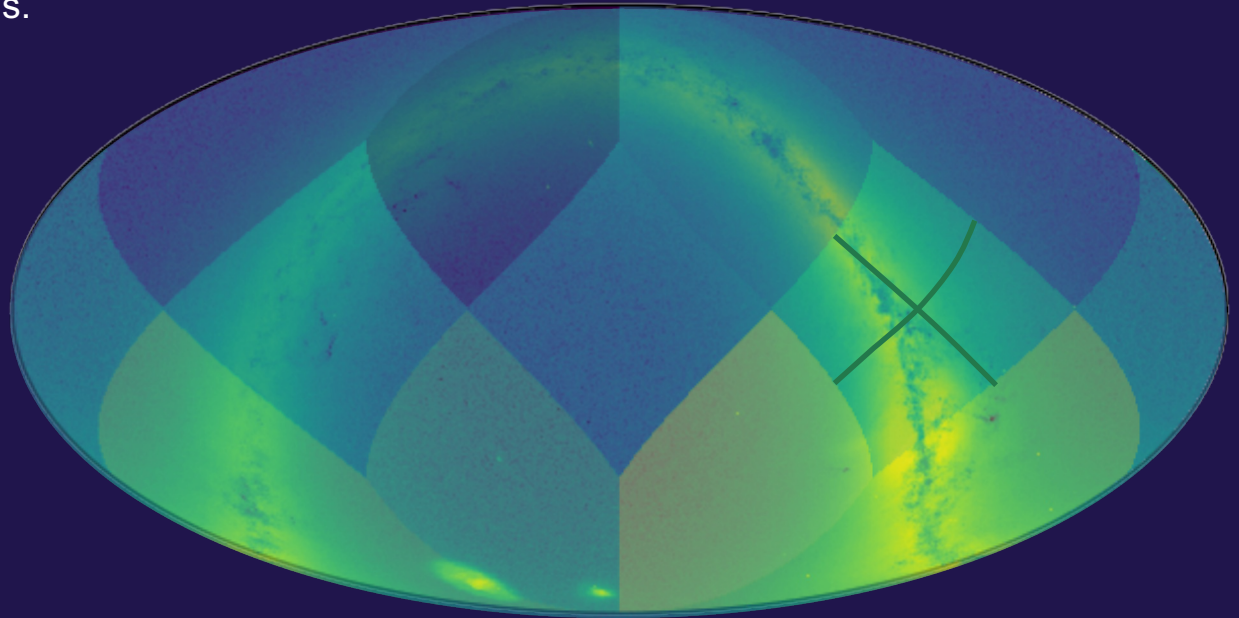


1. Partition Hierarchically

If too many sources fall into a pixel,
split it into four higher order pixels.

Example

```
Norder0-Npix0.csv.gz  
...  
Norder1-Npix28.csv.gz  
Norder1-Npix29.csv.gz  
Norder1-Npix30.csv.gz  
Norder1-Npix31.csv.gz  
...  
Norder0-Npix11.csv.gz
```



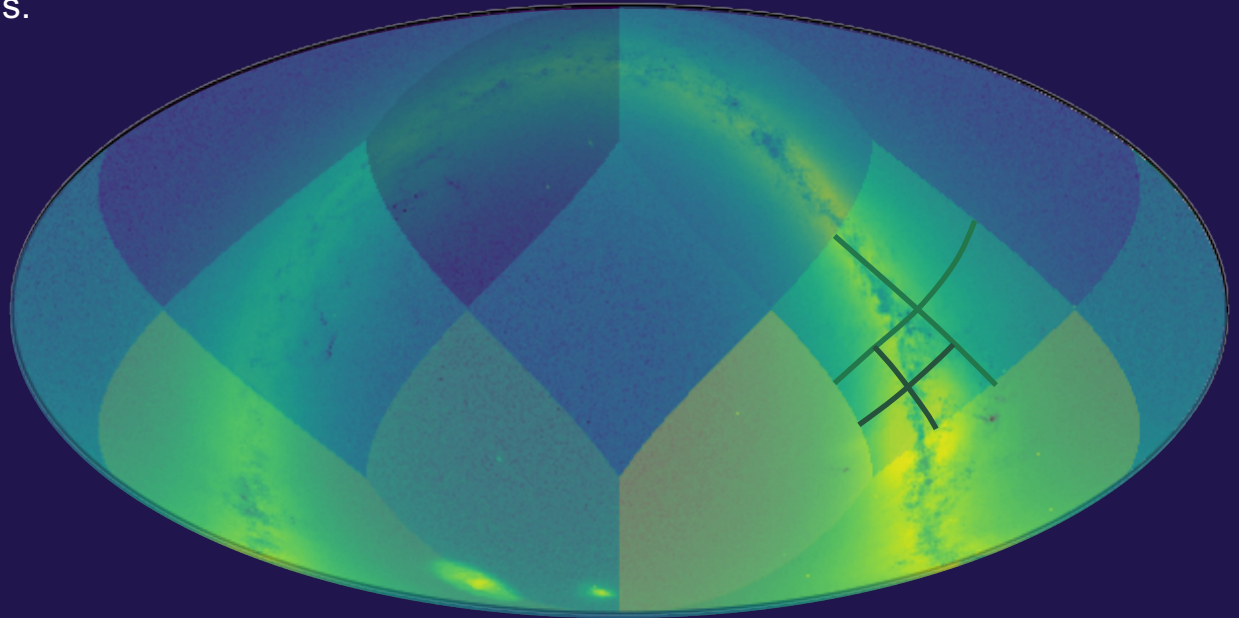
1. Partition Hierarchically

If too many sources fall into a pixel,
split it into four higher order pixels.

Repeat.

Example

```
Norder0-Npix0.csv.gz  
...  
Norder1-Npix28.csv.gz  
Norder1-Npix29.csv.gz  
Norder1-Npix30.csv.gz  
Norder1-Npix31.csv.gz  
...  
Norder0-Npix11.csv.gz
```



1. Partition Hierarchically

If too many sources fall into a pixel,
split it into four higher order pixels.

Repeat.

Example

Norder0-Npix0.csv.gz

...

Norder1-Npix28.csv.gz

Norder1-Npix29.csv.gz

Norder1-Npix30.csv.gz

Norder2-Npix112.csv.gz

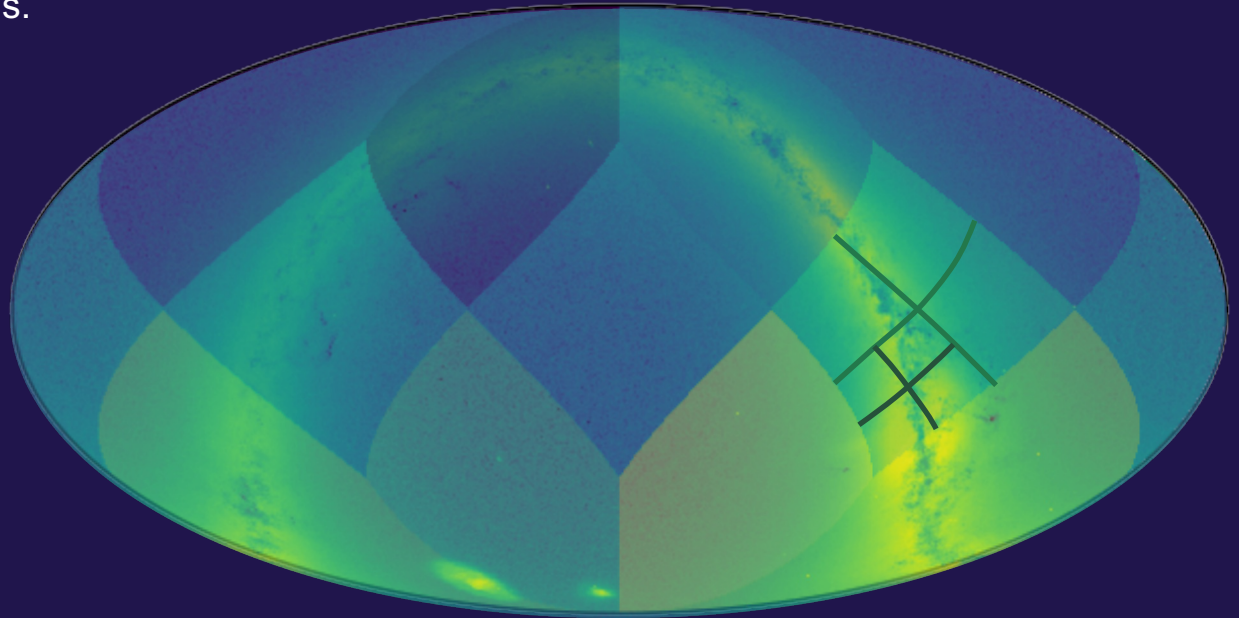
Norder2-Npix113.csv.gz

Norder2-Npix114.csv.gz

Norder2-Npix115.csv.gz

...

Norder0-Npix11.csv.gz



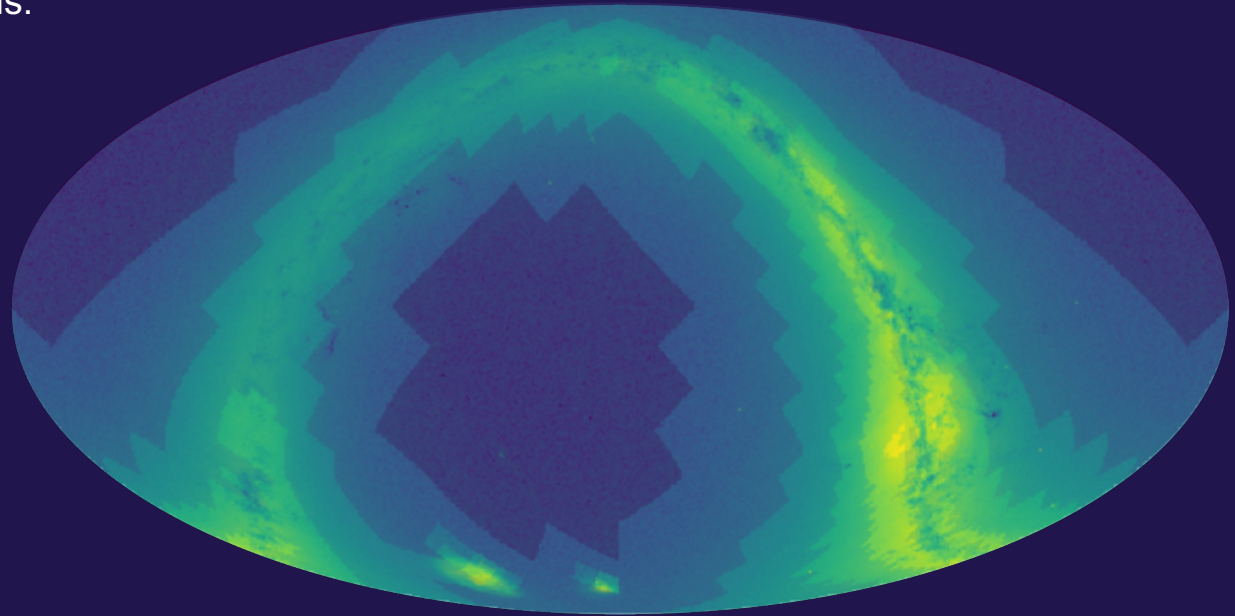
1. Partition Hierarchically

If too many sources fall into a pixel,
split it into four higher order pixels.

Repeat until each file size
is beneath some
pre-defined threshold.

Figure: an overlay of
Gaia counts and the
partitioning map, taking
MAXOBJECTS=1e6

order:



3993 partitions (tiles) for Gaia DR3, with 1M
object/partition threshold

2. Serialization: Parquet in HiPS-like Directory Structure

Store the partition files (tiles) into a directory tree.

Parquet files in the leaf nodes.

```
Norder=0/Dir=0/Npix=0.parquet
```

```
...
```

```
Norder=1/Dir=0/Npix=28.parquet
```

```
Norder=1/Dir=0/Npix=29.parquet
```

```
Norder=1/Dir=0/Npix=30.parquet
```

```
Norder=2/Dir=0/Npix=112.parquet
```

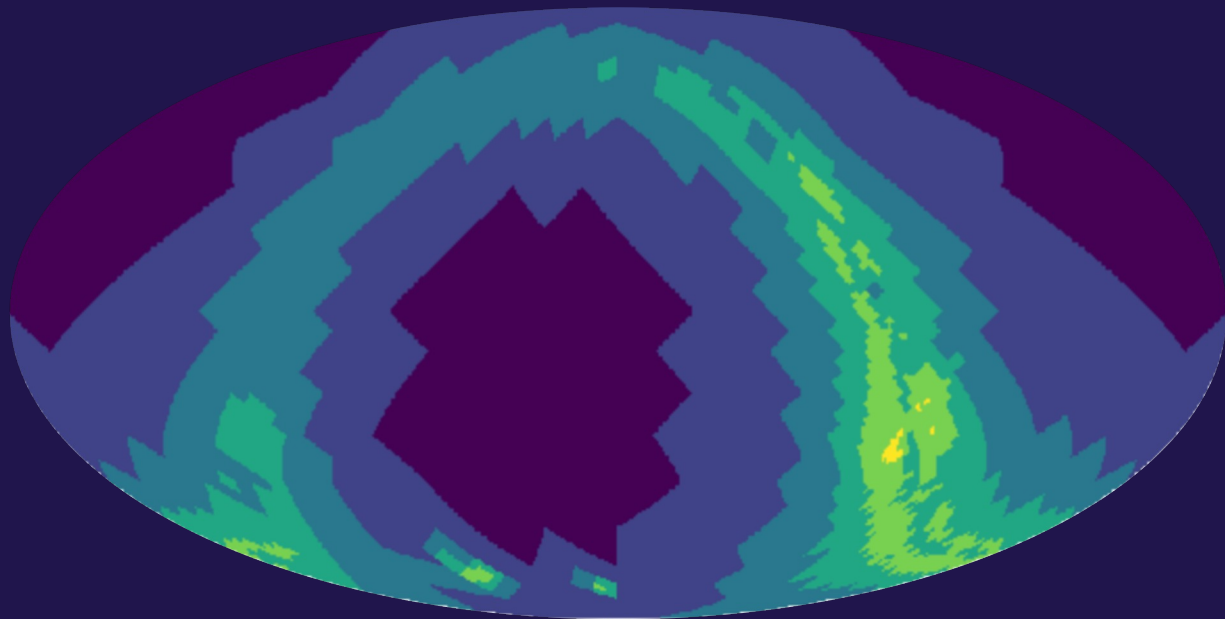
```
Norder=2/Dir=0/Npix=113.parquet
```

```
Norder=2/Dir=0/Npix=114.parquet
```

```
Norder=2/Dir=0/Npix=115.parquet
```

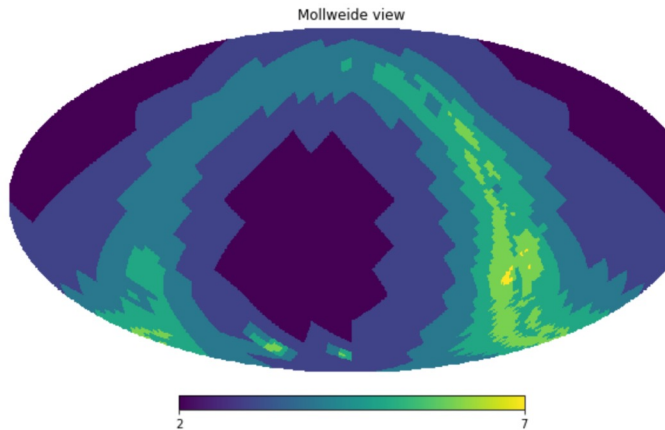
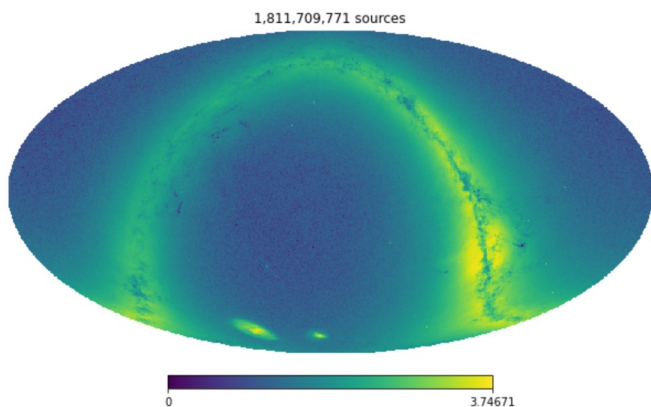
```
...
```

```
Norder=0/Dir=0/Npix=11.parquet
```



All together: Hierarchical Adaptive Tiling Scheme (HATS)

Layout on "disk":



```
gaia/dataset/_metadata
gaia/dataset/_common_metadata
gaia/properties
gaia/partition_info.csv
gaia/point_map.fits
gaia/dataset
gaia/dataset/Norder=2
gaia/dataset/Norder=2/Dir=0
gaia/dataset/Norder=2/Dir=0/Npix=20.parquet
gaia/dataset/Norder=2/Dir=0/Npix=22.parquet
gaia/dataset/Norder=2/Dir=0/Npix=21.parquet
gaia/dataset/Norder=2/Dir=0/Npix=23.parquet
gaia/dataset/Norder=2/Dir=0/Npix=17.parquet
gaia/dataset/Norder=2/Dir=0/Npix=0.parquet
gaia/dataset/Norder=2/Dir=0/Npix=19.parquet
gaia/dataset/Norder=2/Dir=0/Npix=2.parquet
gaia/dataset/Norder=2/Dir=0/Npix=28.parquet
gaia/dataset/Norder=2/Dir=0/Npix=29.parquet
gaia/dataset/Norder=2/Dir=0/Npix=34.parquet
gaia/dataset/Norder=2/Dir=0/Npix=40.parquet
gaia/dataset/Norder=2/Dir=0/Npix=35.parquet
gaia/dataset/Norder=2/Dir=0/Npix=42.parquet
gaia/dataset/Norder=2/Dir=0/Npix=41.parquet
gaia/dataset/Norder=2/Dir=0/Npix=43.parquet
gaia/dataset/Norder=2/Dir=0/Npix=44.parquet
gaia/dataset/Norder=2/Dir=0/Npix=46.parquet
gaia/dataset/Norder=2/Dir=0/Npix=64.parquet
gaia/dataset/Norder=2/Dir=0/Npix=65.parquet
```

Gaia DR2 Catalog Counts (log scale)

*Visualization of file storage (color = healpix level)
3933 partitions of similar size (128-256 MB)*



Updates since Sydney

Format Updates

- Renaming: HiPSCat -> HATS
- Harmonize metadata file naming/contents conventions with HiPS
- Healpix index made non-unique
- Experimental append support
- VOTable metadata in Parquet

```
gaia/dataset/_metadata
gaia/dataset/_common_metadata
gaia/properties
gaia/partition_info.csv
gaia/point_map.fits
gaia/dataset
gaia/dataset/Norder=2
gaia/dataset/Norder=2/Dir=0
gaia/dataset/Norder=2/Dir=0/Npix=20.parquet
gaia/dataset/Norder=2/Dir=0/Npix=22.parquet
gaia/dataset/Norder=2/Dir=0/Npix=21.parquet
gaia/dataset/Norder=2/Dir=0/Npix=23.parquet
gaia/dataset/Norder=2/Dir=0/Npix=17.parquet
gaia/dataset/Norder=2/Dir=0/Npix=0.parquet
gaia/dataset/Norder=2/Dir=0/Npix=2.parquet
gaia/dataset/Norder=2/Dir=0/Npix=19.parquet
gaia/dataset/Norder=2/Dir=0/Npix=28.parquet
gaia/dataset/Norder=2/Dir=0/Npix=29.parquet
gaia/dataset/Norder=2/Dir=0/Npix=34.parquet
gaia/dataset/Norder=2/Dir=0/Npix=40.parquet
gaia/dataset/Norder=2/Dir=0/Npix=35.parquet
gaia/dataset/Norder=2/Dir=0/Npix=42.parquet
gaia/dataset/Norder=2/Dir=0/Npix=41.parquet
gaia/dataset/Norder=2/Dir=0/Npix=43.parquet
gaia/dataset/Norder=2/Dir=0/Npix=44.parquet
gaia/dataset/Norder=2/Dir=0/Npix=46.parquet
gaia/dataset/Norder=2/Dir=0/Npix=64.parquet
gaia/dataset/Norder=2/Dir=0/Npix=65.parquet
gaia/dataset/Norder=2/Dir=0/Npix=70.parquet
gaia/dataset/Norder=2/Dir=0/Npix=67.parquet
gaia/dataset/Norder=2/Dir=0/Npix=72.parquet
```

Code on GitHub/PyPI/conda-forge



Astronomy Data Commons

Software Infrastructure for Science Platforms and Scalable Astronomy on Cloud Resources

Pinned

[Customize pins](#)

axs Public

Astronomy eXtensions for Spark: Fast, Scalable, Analytics of Billion+ row catalogs

Python 23 12

hats Public

Hierarchical Adaptive Tiling Scheme

Python 17 5

Low level format routines

lsdb Public

Large Survey DataBase

Python 19 5

End-user analytics tool

hats-import Public

HATS import - generate HATS-partitioned catalogs

Python 6 5

Robust importer

<https://github.com/astronomy-commons>

All of these are also available on PyPI and conda-forge.



LSDB: Python Analytics for HATS

- LSDB: Large Survey Database
- Enable Pandas-like analysis on trillions of observations with thousands of cores
- Build on existing tools: Dask (looking at Ray).
- Full HATS awareness: spatial queries, cross-matching, timeseries, multi-dataset joining.
- Stabilizing the API

```
img = gaia
    .query("pm > 10")
    .crossmatch(ztf)
    .join(ztf_sources)
    .for_each(varstar_classify)
    .query("pRRLy > 0.95")
    .skymap()

hp.mollview(img)
```

LSDB target APIs: The API center science. Multi-processing, autoscaling, fail-over, etc. are all implicit. Good user experience.

Wyatt et al. (2023)

<https://github.com/astronomy-commons/lldb>

Nested Pandas: Compact/lazy JOIN evaluation

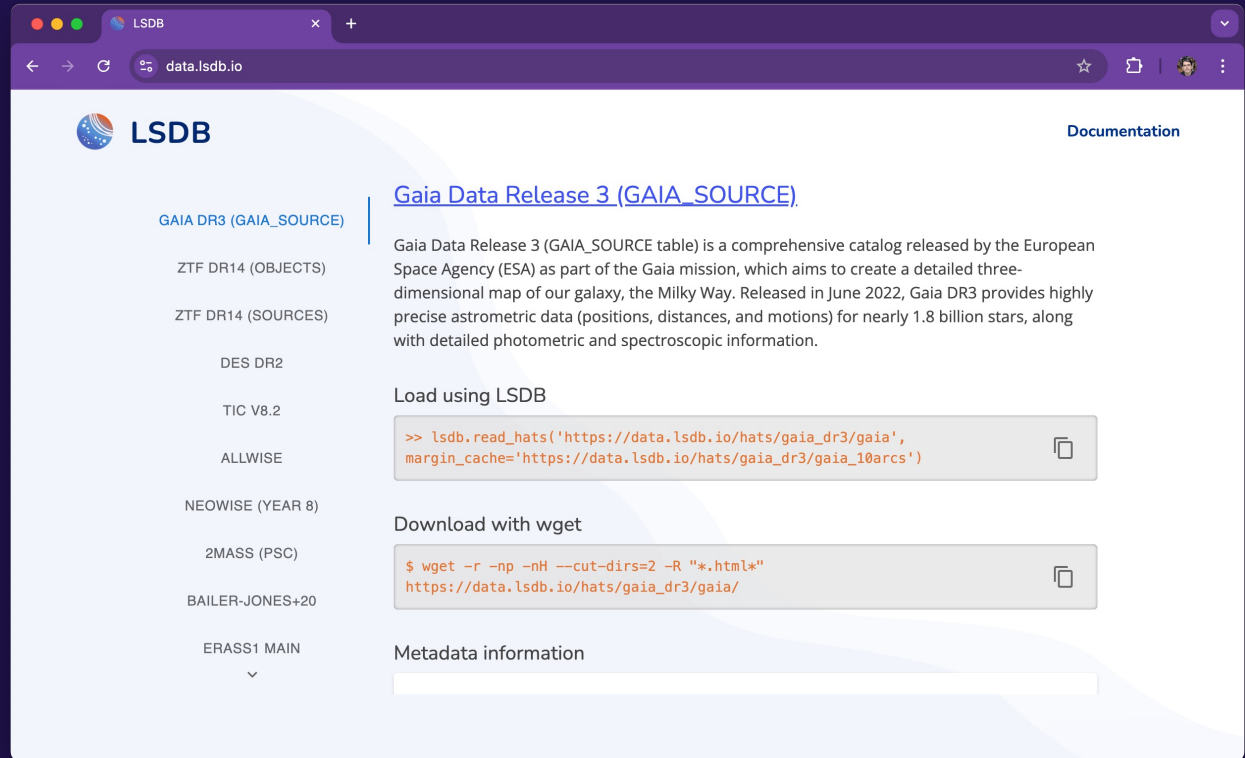
- Typical catalogs are structured as Object tables, to be joined to Source/Detection tables (time series).
- SQL JOIN-ing Object to Source tends to duplicate fields from Object, and generates a flat table that's inconvenient for the end-user to work with

	a	b	nested
0	0.417022	0.184677	t flux band 0 8.383890 1...
1	0.720324	0.372520	t flux band 0 13.704390 1...
2	0.000114	0.691121	t flux band 0 4.089045 9...
3	0.302333	0.793535	t flux band 0 17.562349 6...
4	0.146756	1.077633	t flux band 0 0.547752 ...

- nested_pandas is a memory-efficient pandas extension that presents a “time-series” view to the user, while retaining the speed of evaluation equivalent to that of a flat table.

HATS Datasets @ <https://data.lsdب.io>

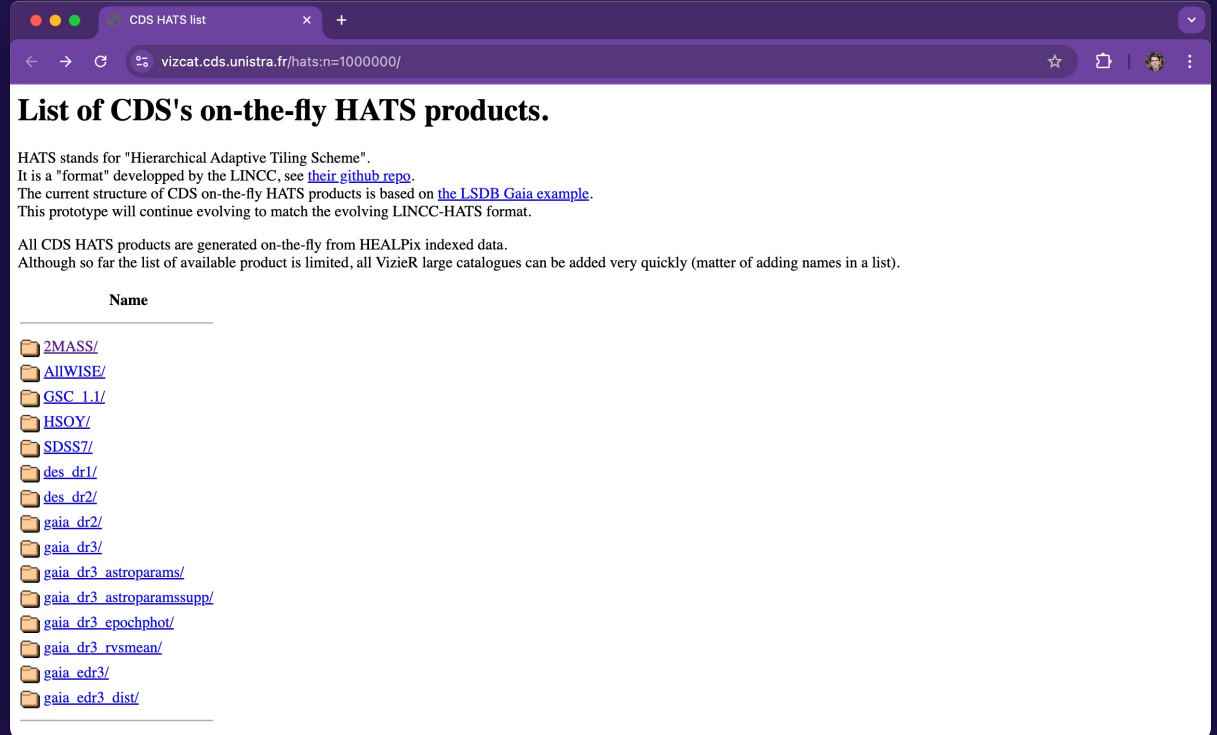
- A library of HATS-partitioned datasets to support testing and experimentation with the format
- Download LSDB (the main HATS analytics tool), point it to this website, and immediately start running queries.



The screenshot shows the LSDB website interface. The browser address bar displays `data.lsdب.io`. The page header includes the LSDB logo and a "Documentation" link. A sidebar on the left lists various datasets: GAIA DR3 (GAIA_SOURCE), ZTF DR14 (OBJECTS), ZTF DR14 (SOURCES), DES DR2, TIC V8.2, ALLWISE, NEOWISE (YEAR 8), 2MASS (PSC), BAILER-JONES+20, and ERASS1 MAIN. The main content area is titled "Gaia Data Release 3 (GAIA_SOURCE)" and contains a descriptive paragraph about the dataset. Below the description, there are two code blocks: one for loading the data using LSDB and another for downloading it with wget. The LSDB code block shows the command `>> lsdb.read_hats('https://data.lsdب.io/hats/gaia_dr3/gaia', margin_cache='https://data.lsdب.io/hats/gaia_dr3/gaia_10arcs')`. The wget code block shows the command `$ wget -r -np -nH --cut-dirs=2 -R "*.html*" https://data.lsdب.io/hats/gaia_dr3/gaia/`. At the bottom, there is a section for "Metadata information" with a placeholder box.

HATS Facade @ CDS

- An experiment by FX Pineau et al. to test if CDS catalog holdings can be on-the-fly represented as HATS.
- Converts internal CDS database representation of the catalogs into HATS-partitioned files, served over HTTP.



List of CDS's on-the-fly HATS products.

HATS stands for "Hierarchical Adaptive Tiling Scheme".
It is a "format" developed by the LINCC, see [their github repo](#).
The current structure of CDS on-the-fly HATS products is based on [the LSDB Gaia example](#).
This prototype will continue evolving to match the evolving LINCC-HATS format.

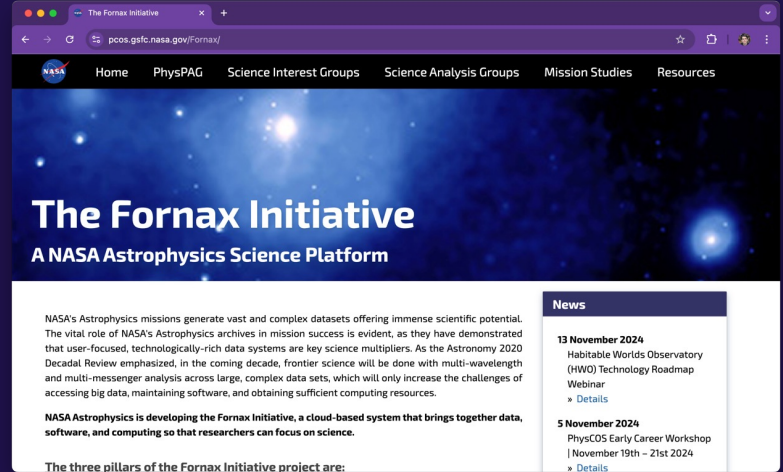
All CDS HATS products are generated on-the-fly from HEALPix indexed data.
Although so far the list of available product is limited, all Vizier large catalogues can be added very quickly (matter of adding names in a list).

Name

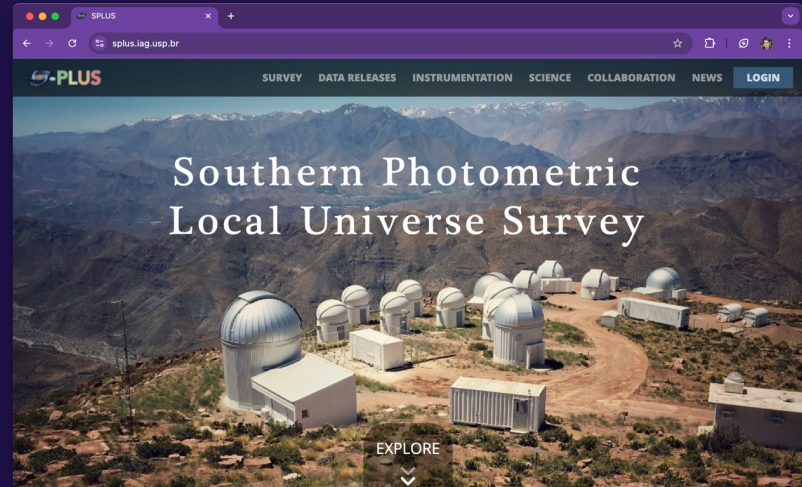
- [2MASS/](#)
- [AllWISE/](#)
- [GSC 1.1/](#)
- [HSOY/](#)
- [SDSS7/](#)
- [des_dr1/](#)
- [des_dr2/](#)
- [gaia_dr2/](#)
- [gaia_dr3/](#)
- [gaia_dr3_astroparams/](#)
- [gaia_dr3_astroparamssupp/](#)
- [gaia_dr3_epochphot/](#)
- [gaia_dr3_rvsmean/](#)
- [gaia_cdr3/](#)
- [gaia_cdr3_dist/](#)

Fornax, SPLUS, ...

- Continuing the integration and testing of HATS/LSDB within Fornax (IPAC & STScI Fornax teams)
- A “HATS/Parquet” server for SPLUS with server-size predicate evaluation (Gustavo Schwarz, Mackenzie U.)



The screenshot shows the NASA Fornax Initiative website. The browser address bar displays "pcos.gsfc.nasa.gov/Fornax/". The navigation menu includes "Home", "PhysPAG", "Science Interest Groups", "Science Analysis Groups", "Mission Studies", and "Resources". The main heading is "The Fornax Initiative" with the subtitle "A NASA Astrophysics Science Platform". The page content includes a paragraph about NASA's Astrophysics missions generating vast datasets, a "News" sidebar with two entries: "13 November 2024 Habitable Worlds Observatory (HWO) Technology Roadmap Webinar" and "5 November 2024 PhysCOS Early Career Workshop | November 19th - 21st 2024", and a section titled "The three pillars of the Fornax Initiative project are:".



The screenshot shows the SPLUS website. The browser address bar displays "splus.lag.usp.br". The navigation menu includes "SURVEY", "DATA RELEASES", "INSTRUMENTATION", "SCIENCE", "COLLABORATION", "NEWS", and "LOGIN". The main heading is "Southern Photometric Local Universe Survey" over a background image of a mountain range with several observatory domes. A button labeled "EXPLORE" is visible at the bottom.

Status and Looking Ahead

- Format is in a reasonably solid draft form now
 - Remains to be done: optimized storage of distributed cross-match margins
- Working or in touch with most major archives/projects.
- Preparing an IVOA Note.

- Application support is maturing for Python (LSDB), but we're looking to expand support in other ecosystems (Rust, Java)
 - E.g. a Spark DataSource for reading HATS would enable advanced support for HATS within Spark
- Field testing within Rubin Commissioning
- Thinking about Iceberg serialization

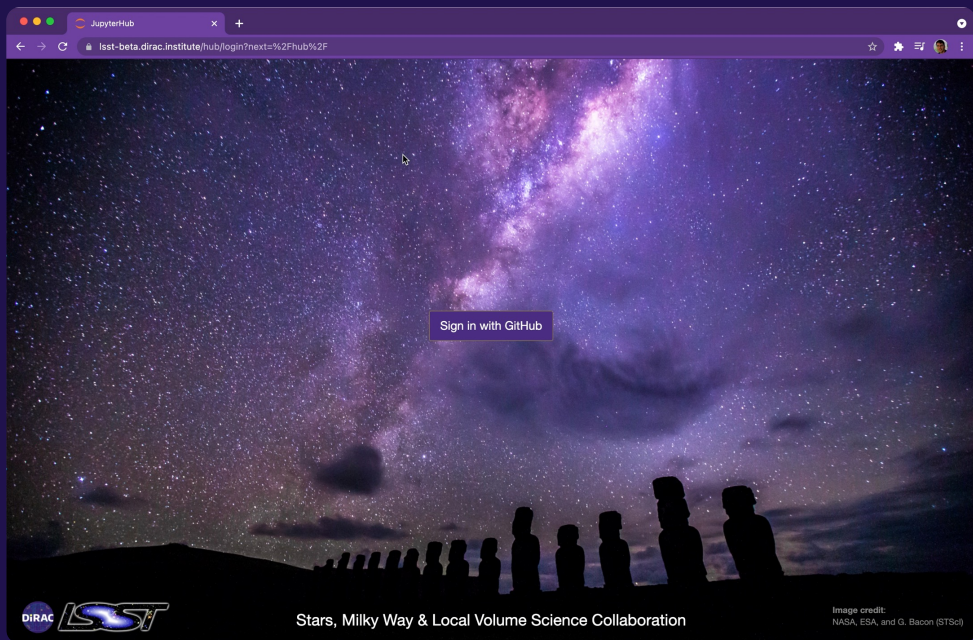
Where to Try



LINCC Hub: <https://lsst.dirac.dev>

- LINCC Hub: *The LINCC Science Platform for the LSST Science Collaborations* (email ncaplar@uw.edu for access)

- NASA Astrophysics Science Platform: Working with the Fornax team at IRSA/MAST/HEASARC to deploy and test.



- Or run `conda install -c conda-forge lsdb` on your fav. science platform.

Summary and next steps

Mailing list: <https://groups.google.com/g/hats-wg>
Repositories: <https://github.com/astronomy-commons>
Meetings: 10am PT, every third Friday of the month

- Aiming to enable end-user analyses on 1-100T+ catalog datasets. Ad-hoc collaboration of scientists/engineers from LINCC, Rubin, MAST, IRSA, HESARC, LINeA, CDS. Developing formats and tools.
- Lots of progress since last year. Data format and Python tools solidifying. We've gathered enough initial real-world usage to start drafting the formal format spec. Preparing an IVOA Note.
- Next year:
 - Perfect the tools (UX and performance are a priority)
 - Larger scale user deployment
 - Would love to get add'l feedback and implementations (e.g. Java or Rust)



Collaboratively advancing data-intensive astronomy.



AST-2003196

Thank You !

Contact: mjuric@uw.edu

UNIVERSITY of WASHINGTON