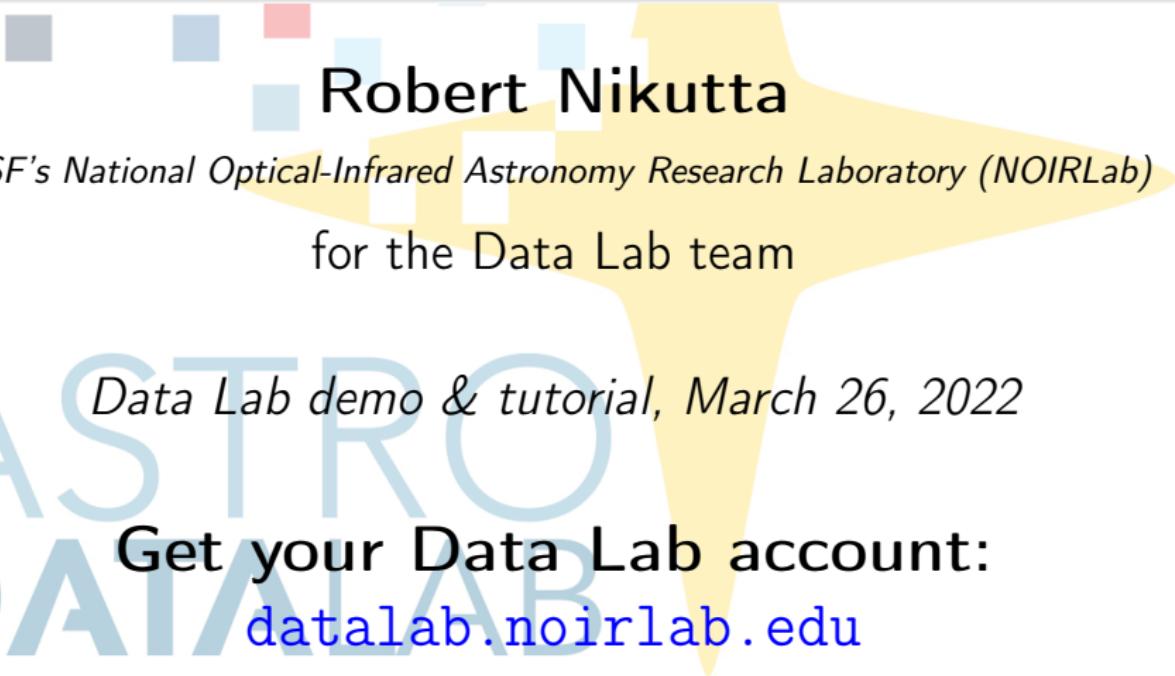


Astro Data Lab

Open-Data, Open-Access Science Platform



Robert Nikutta

NSF's National Optical-Infrared Astronomy Research Laboratory (NOIRLab)

for the Data Lab team

Data Lab demo & tutorial, March 26, 2022

Get your Data Lab account:
datalab.noirlab.edu

What is Data Lab today?

Soft side:

- ▶ Community of now ~2000 users (*including you!*)
- ▶ Team of scientists, developers, data analyst, DB architects, engineers, sysads (*make it work, enhance datasets, engage with users a lot*)

Hard side:

- ▶ Repository for large (photometric) datasets (*catalogs & images*)
- ▶ New: spectroscopic data : SDSS, BOSS, eBOSS: and soon: DESI EDR
- ▶ Services to access data (*VO, SQL/ADQL, TAP, SIA, ...*)
- ▶ Gateway to other data centers
- ▶ Website / authenticated portal
- ▶ Exploratory tools (*survey coverage, catalog overlay, ...*)
- ▶ Cross-matching service
- ▶ Visualization (*all of Python, APIs*)
- ▶ Analysis facilities (*Jupyter notebook server, common-task helpers*)
- ▶ Remote storage (*vospace, mydb, notebook space*)

In the near future

- ▶ Science platform also for large spectroscopic data (e.g. DESI)
- ▶ Batch compute (*job manager*)
- ▶ Going gradually to the cloud (better on-demand scalability, growth)
- ▶ Survey data publication facility (*BYOD*)
- ▶ Machine learning (*e.g. simple classifiers*)

Data Lab is a:

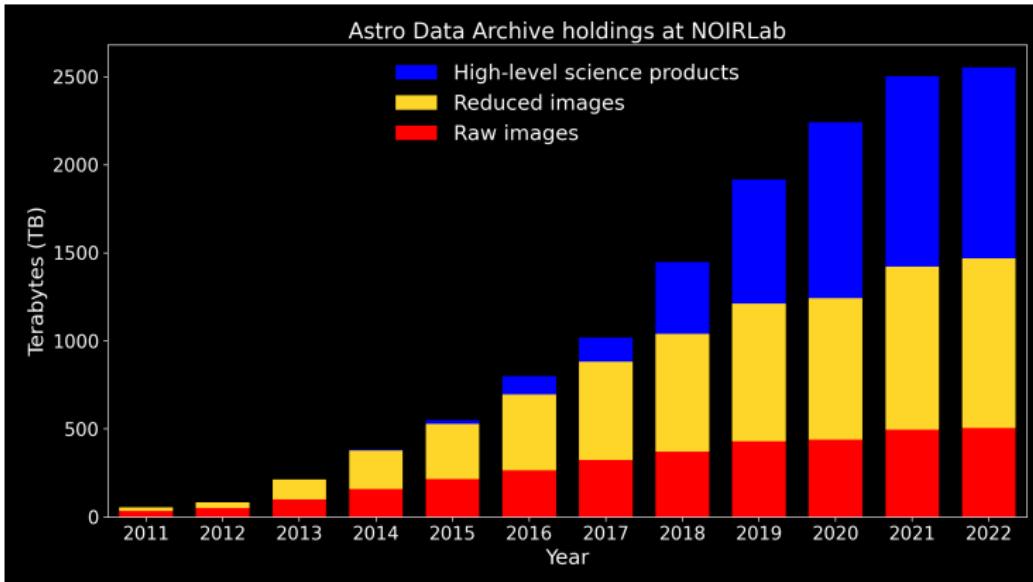
- ▶ Science Platform
- ▶ Data Exploration & Analysis environment
- ▶ Archive 2.0

(Pick your favorite fashionable word)

Why Data Lab?

- ▶ *Originally*: to exploit NOAO (now NOIRLab) public data (4-m telescopes), surveys, time-domain; *specifically*: DES
- ▶ Bring user's analyses to the data
- ▶ Joint analysis with other datasets (ingested by us, or uploaded by user)
- ▶ Preparation of users for PB-scale data now, readiness for LSST
- ▶ *Today*: DES just one of 17 major surveys hosted at Data Lab

Some numbers on data holdings



- ▶ At Data Lab: 130 TB of catalog data, 220+ billion rows, plus 78B rows in crossmatch tables, plus 12B rows in image search tables
- ▶ Images, masks, coadds, etc.: close to 5 PB (at ADA & Data Lab)

Some hosted datasets

A selection:

- ▶ DES DR1, DR2 (691e6 objects)
- ▶ DECaLS DR 3, 4, 5, 6, 7, 8, 9 (1.9e9 objects)
- ▶ DECaPS DR1 (2e9 objects)
- ▶ NOIRLab Source Catalog DR1, DR2 (3.9e9 objects, 68e9 measurements)
- ▶ Gaia DR1, DR2, EDR3 (1.8e9 objects)
- ▶ PHAT (7.5e9 objects)
- ▶ AllWISE (750e6 objects)
- ▶ unWISE (2.2e9 objects)
- ▶ SMASH DR1, DR2 (360e6 objects)
- ▶ SDSS DR 13, 14, 16 (5.1e6 objects)
- ▶ Stripe82 pre-cross-matched
- ▶ LSST simulated catalogs MW + LMC + SMC (>10e9 objects)
- ▶ CatWISE2020 (1.9e9 objects)
- ▶ S-PLUS DR1, DR2 (35e6 together)
- ▶ SkyMapper DR1, DR2 (500e6 objects, 5e9 measurements)
- ▶ UKIDSS DR11plus (1.2e9 objects)
- ▶ VHS DR5 (1.4e9 objects)
- ▶ DELVE DR1, DR2 (445e6 objects)

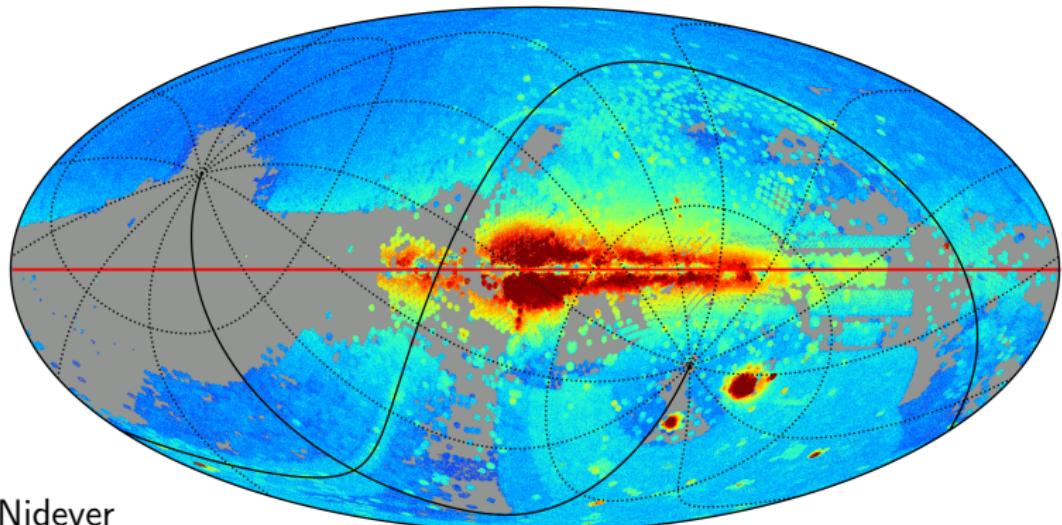
Temporal information in: NSC, SMASH, LS DR 8+9, DELVE DR2,
CatWISE2020, ...



Contact: robert.nikutta@noirlab.edu

NOIRLab Source Catalog

Log number of objects / deg² in the NSC



PI: D. Nidever

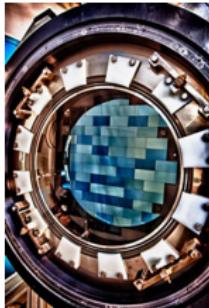
- ▶ DR2: 3.9e9 sources, >68e9 measurements

Some future datasets

- ▶ DESI spectra (Dark Energy Spectroscopic Instrument; 20M galaxies with precise redshifts)
- ▶ LS DR10
- ▶ DELVE DR3
- ▶ Gaia DR3
- ▶ “light” LSST
- ▶ “light” Pan-STARRS
- ▶ SDSS DR17
- ▶ some X-ray catalogs
- ▶ Your data? (data publication service)

Next-gen wide-field cameras

Dark Energy Camera



570 Mpix

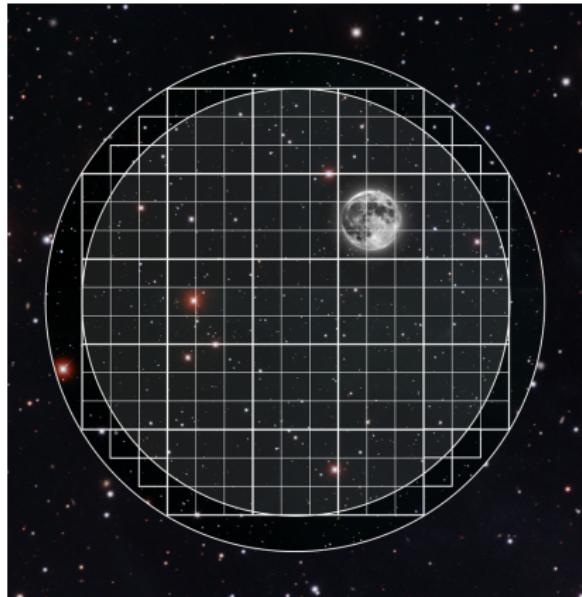
Rubin/LSST Camera



3200 Mpix

Next-gen wide-field cameras

Rubin/LSST focal plane



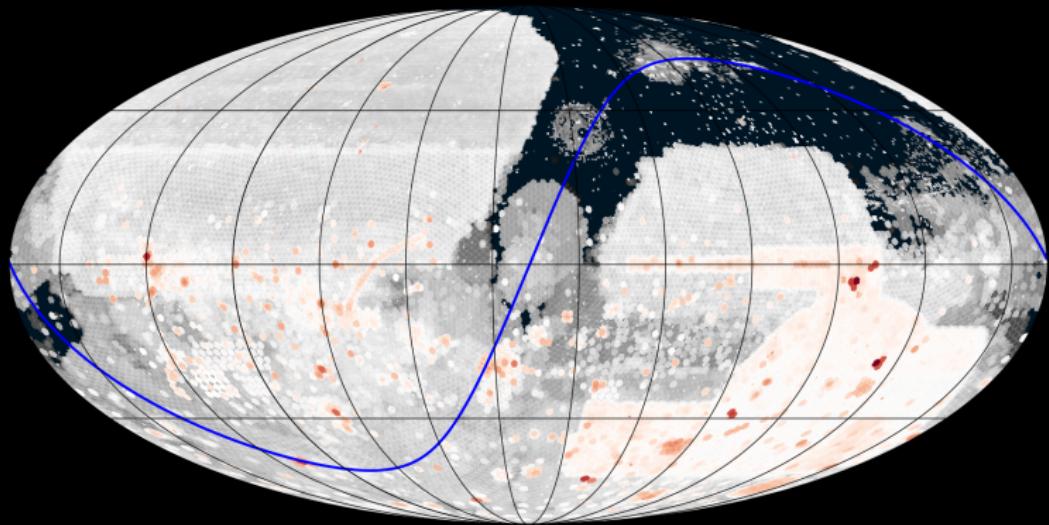
40 x Moon area



Contact: robert.nikutta@noirlab.edu

Sky coverage over time

2019.9



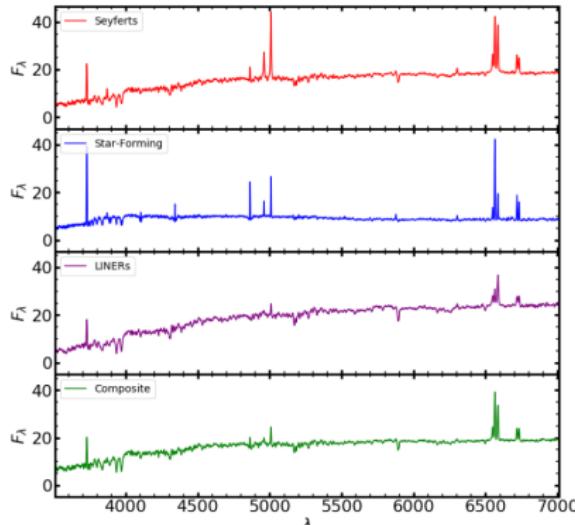
Exposure time at 2-4m telescopes in NSF's OIR Lab Science Data Archive



Contact: robert.nikutta@noirlab.edu

Spectroscopy services

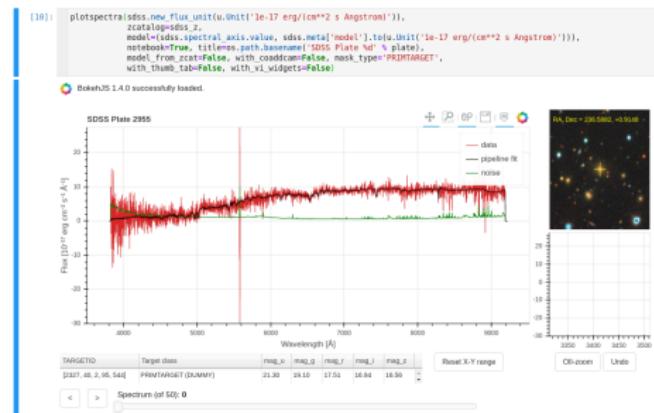
- ▶ Data Lab now has a fast spectroscopic service in production
- ▶ Demonstrator, serving SDSS DR8-16
- ▶ 100+ times faster than opening FITS files
- ▶ Testbed and prototype for new development of a scalable service ready for DESI (codename: SPARCL)



Spectro viewer / inspection tool

We needed a spectral viewer / inspection tool

- ▶ Made DESI's prospect tool stand-alone
- ▶ Based on Bokeh
- ▶ Adapted it to work with SDSS spectra
- ▶ Takes specutils objects



- ▶ Supports pan/zoom, line identification, redshifting, template fitting, thumbnails
- ▶ Goal: make it a generic tool for any future spectro datasets
- ▶ Currently unclear: how to handle thousands and millions of spectra

Current functionality

What	How
Authentication	web, datalab cmd line tool, authClient.py; <i>anonymous user exists</i>
Sky exploration	survey discovery tool (Aladin lite), catalog overlay (web)
Catalog query	web site, queryClient.py, datalab, TAP clients (e.g. TOPCAT)
Image query	Simple Image Access via website or notebooks (scriptable)
Spectro query	Python API via notebooks or scripts (fast!)
Query results...	to VOSpace, to mydb, or return to client
File transfer	datalab, storeClient.py
File services	collections of survey data, including suppl. files
Cross-matching	web tool, APIs
Visualization	notebooks, APIs
Analysis	Jupyter notebooks (all relevant packages present), or do locally
Common tasks	in d1.helpers.* (e.g. convert to pandas, sky plot, ...)

User = someone with a DL account

Register as a new user

Username:

Must use between 6 and 30 characters, only lower case letters, numbers and underscores allowed.

Password:

Must use 6 or more characters, and must contain UPPER/lowercase and numbers.

Full Name:

Email:

Privacy: Your email address will not be shared or sold to third parties.

Affiliation:

Anti-spam verification:

I'm not a robot

reCAPTCHA
Privacy - Terms

[Register](#)

You...

- ▶ Get 1 TB of storage on vospace
- ▶ Get 250 GB of MyDB storage
- ▶ Can upload datasets for joint analysis
- ▶ Can edit/create/delete notebooks
- ▶ Can upload own Python source code
- ▶ Can share data with others
- ▶ Your workflows are persistent

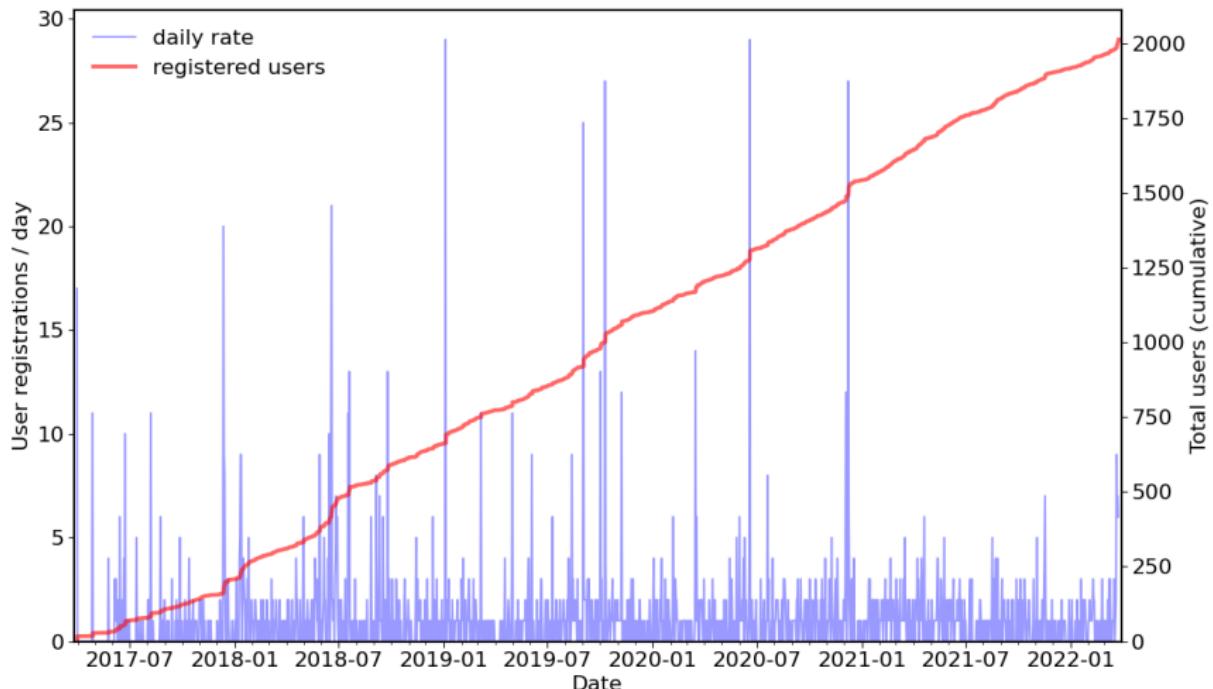
Get your account today

Go to: datalab.noirlab.edu (click on 'Sign up')



Contact: robert.nikutta@noirlab.edu

User community



Contact: robert.nikutta@noirlab.edu

Curated default notebooks

astro-datalab / notebooks-latest Public

Code Issues 6 Pull requests Discussions Actions Projects Wiki Security Insights Settings

master 11 branches 0 tags Go to file Add file Code

 rmikutta	Merge pull request #122 from DavidVargasMora/LarsChangesMar22	... f36b87c 5 days ago	473 commits
 01_GettingStartedWithDataLab	noaodatalab --> astro-datalab	last month	
 02_DataAccessOverview	- noao to noirlab NB update batch 1 (both ipynb and html files)	4 months ago	
 03_ScienceExamples	noaodatalab --> astro-datalab	last month	
 04_HowTos	noaodatalab --> astro-datalab	last month	
 05_Contrib	Upload ANTARES "RCB Stars" HTML	last month	
 06_EPO	Resolved merge conflicts	6 days ago	
 tests	- Exclude e-TeenAstronomyCafe NBs from default testing.	10 months ago	
 .gitignore	Set lineWidth parameter to 3, on lightcurve_slider.py in order to solv...	6 months ago	
 CONTRIBUTING	Update README.txt and CONTRIBUTING with new 'astro-datalab' ...	last month	
 DataLabNotebookTemplate.ipynb	noaodatalab --> astro-datalab	last month	
 LICENSE	- noao to noirlab NB update batch 1 (both ipynb and html files)	4 months ago	
 README.txt	Update README.txt and CONTRIBUTING with new 'astro-datalab' ...	last month	
 keywords.txt	Sort keywords.txt	3 years ago	
 test_ignore.txt	test15.ignore	3 months ago	

About

Default set of Data Lab notebooks, by DL team and contributed by users

Readme

BSD-3-Clause License

17 stars

12 watching

26 forks

Releases

No releases published

Create a new release

Packages

No packages published

Publish your first package

Contributors 22



Intros, How-Tos, Data Access, complete Science Cases, ANTARES filter dev kit, EPO, ...



Contact: robert.nikutta@noirlab.edu

Helpdesk & User Manual online

The screenshot shows a web browser window titled "Data Lab Help Desk" with the sub-page "NOAO Data Lab User Forum". The main content area features a large globe map with blue and green data points. Overlaid on the globe are three callout boxes: "Search answers for all your queries" (with a magnifying glass icon), "One destination for all your queries" (with a question mark icon), and "Get answers from the experts" (with a checkmark icon). Below the globe, there's a section titled "Recent questions and answers" with a search bar and a button "Ask a question:". A specific post is highlighted: "Unable to login to Detalab" by "klausen" (300 points) with 68 views. The bottom of the page has footer links for "Data Lab Home", "File Edit View Tools Window Help", and "Data Lab Help Desk".

Ask a question (also anonymously),
FAQ, humans take care of you



AURA



Contact: robert.nikutta@noirlab.edu

The screenshot shows a web browser window titled "NOAO Data Lab 1.1.0 documentation". The main content area has a header "Welcome to the NOAO Data Lab documentation". Below it is a "Table Of Contents" section with a yellow starburst logo. The table of contents lists various topics: 1. Using the NOAO Data Lab, 2. FAQs, 3. Appendices, 4. Data Publication docs, and Indices and tables. Under "Indices and tables", there are links for "Index", "Module Index", and "Search Page". At the bottom, there's a footer with "NOAO Data Lab 1.1.0 documentation", copyright information ("© Copyright 2016-2017, NOAO Data Lab <detalab@noao.edu>"), and a note ("Last updated on 2018-Jan-07").

Overviews, glossary, science cases,
SQL examples, tips & tricks

API documentation

File Edit View Tools Window Help documentation (auto-generated) --- datalab - next | modules | index

datalab 1.1.0 documentation »

Table of Contents

NOAO Data Lab API documentation (auto-generated)
Indices and tables

Next topic
authClient module

This Page
Show Source

Quick search

Go

NOAO Data Lab API documentation (auto-generated)

Contents:

- authClient module
- queryClient module
- storeClient module
- helpers module
 - cluster helpers
 - crossmatch helpers
 - plot helpers
 - utils helpers
 - legacy helpers
 - all helpers
- dlinterface module
- dltasks module

Indices and tables

- Index
- Module Index
- Search Page

next | modules | index

© Copyright 2017, DataLab team. Last updated on 2017-Dec-20. Created using Sphinx 1.6.3.

File Edit View Tools Window Help Client module --- datalab 1.1.0 documentation previous | next | modules | index

datalab 1.1.0 documentation »

authClient module

Previous topic
NOAO Data Lab API documentation (auto-generated)

Next topic
queryClient module

This Page
Show Source

Quick search

Go

authClient

Bases: `object`

AUTHCLIENT – Client-side methods to access the Data Lab Authentication Service.

`debug(debug_val)` [source]

`get_profile()` [source]

Get the requested service profile.

Parameters: `None` –
Returns: `profile` – The currently requested service profile.
Return type: `str`

Example

```
from dl import authMgr
profile = authMgr.client.get_profile ()
```

`get_service()` [source]

Return the currently-used Authentication Service URL.

Parameters: `None` –
Returns: `service_url` – The currently-used Authentication Service URL.
Return type: `str`

Example

```
from dl import authMgr
service_url = authMgr.client.get_service ()
```

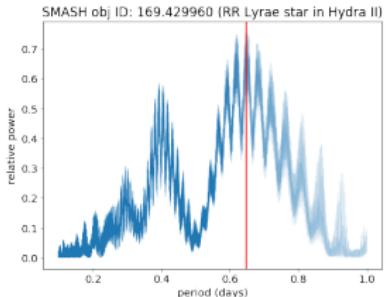
`hasAccess(user, resource)` [source]



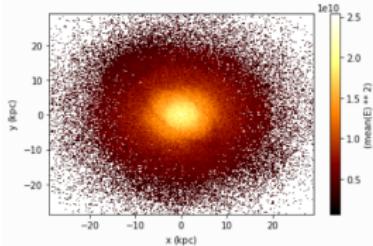
Contact: robert.nikutta@noirlab.edu

Visualization examples

matplotlib



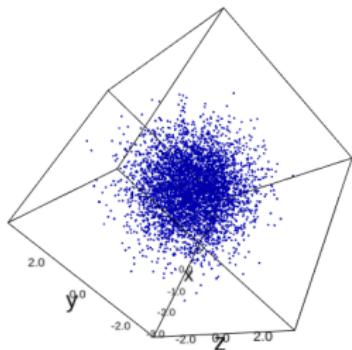
vaex



ipyvolume / ipywidgets

Also

- ▶ bokeh
- ▶ exploring Glue, Firefly
- ▶ spectra visualization tools
(e.g. specviz)
- ▶ ...



Contact: robert.nikutta@noirlab.edu

Contact the Data Lab team

✉ datalab@noirlab.edu

🌐 datalab.noirlab.edu

🐙 github.com/astro-datalab

🐦 [@DataLabAstro](https://twitter.com/DataLabAstro)

Thank you!

ଧ୍ୟାନପଦ୍ଧତିରୁ!

Galatoomaal!

Хвала вам!

robert.nikutta@noirlab.edu



Hacking together on great science with
Data Lab

Helpful links

User manual

<https://datalab.noirlab.edu/docs/manual/>

Help desk

<https://datalab.noirlab.edu/help/>

On the Jupyter notebook server, and Data Lab example/how-to notebooks

<https://datalab.noirlab.edu/docs/manual/UsingTheNOAODataLab/JupyterNotebooks/JupyterNotebooks.html>

How to copy latest Data Lab notebooks to a new (writable) directory

<https://datalab.noirlab.edu/docs/manual/UsingTheNOAODataLab/JupyterNotebooks/JupyterNotebooks.html#get-the-latest-set-of-default-notebooks>

Notebooks on Github (check out README and CONTRIBUTING files)

<https://github.com/astro-datalab/notebooks-latest/>



Contact: robert.nikutta@noirlab.edu