Exploring DESI DR1 efficiently with the **SPARCL** spectral database (and the **Astro Data Lab** science platform)

Value-added services to maximize discovery

Stéphanie Juneau

On behalf of the NSF NOIRLab Astro Data Lab & SPARCL teams

(SPARCL: A. Jacques, P. Peterson, B. Weaver, S. McManus, R. Nikutta, A. Bolton, R. Pucha)

DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

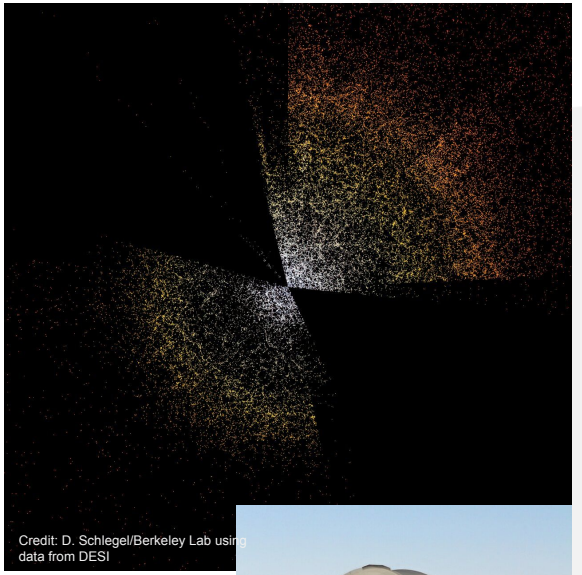
DESI (Dark Energy Spectroscopic Instrument)

General DESI website: www.desi.lbl.gov
DESI Data documentation website: data.desi.lbl.gov/doc/

- Primarily a DOE cosmology experiment (stage IV) but also a treasure trove of data for astronomers and students!

- Data Releases
  - **DESI EDR**: Survey Validation; spectra of 1.7 million unique targets
  - **DESI DR1**: Data up to 13 months (“Year 1”); 23 million spectra of galaxies, quasars & stars
  - Future data releases will keep growing (DR2: “Year 3” data with >60 million)

- DESI at NOIRLab:
  - Data in the Astro Data Archive: astroarchive.noirlab.edu
  - Catalog Database in the Astro Data Lab: datalab.noirlab.edu
  - Spectral Database as part of SPARCL: astrosparcl.datalab.noirlab.edu


Credit: D. Schlegel/Berkeley Lab using data from DESI

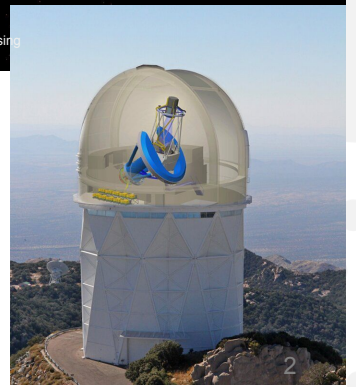


Image credit: R. Lafever, J. Moustakas/DESI Collaboration, P. Marenfeld/NOIRLab/AURA/NSF & E. Acosta/Vera C. Rubin Observatory/ NOIRLab/ NSF/ AURA

SPARCL

SPectra Analysis & Retrievable Catalog Lab

astrosparcl.datalab.noirlab.edu

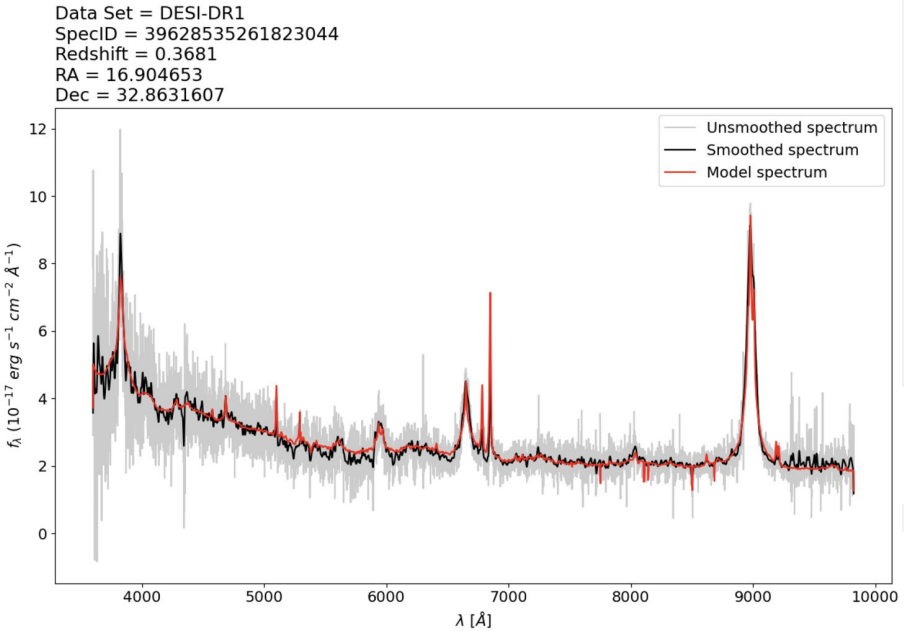
- Spectroscopic database for large surveys/datasets
- Data Discovery
- Data Access/Retrieval
- Server can work with different clients
- Installed at Astro Data Lab (works in JupyterLab)
- Current datasets:
  - **SDSS DR16** (SDSS, BOSS), **DESI EDR** publicly available
  - **DESI DR1** (with authentication)

- Future goals:
  - Prepare DESI DR2 as soon as possible
  - Add other spectroscopic datasets (streamlining the ingest process)
  - Add more advanced functionality (e.g., aligning spectra)

Astronomy friendly format

Value-added data product for the healpix-coadded spectra

- Goal: similar to SDSS 1D spectra data format for ease of use
  - Combined across cameras (single vector for **wavelength**, **flux**, **ivar**, etc.)
  - Rendered best-fitting template from Redrock (**model**)
  - *Still missing:* **sky** spectrum



Currently at NERSC as part of GQP:

`/global/cfs/cdirs/desi/science/gqp/camcoadd/`

- `fuji/healpix/`
- `iron/healpix/`

Notebooks: How to use SPARCL

How to use SPARCL

SPectra Analysis and Retrievable Catalog Lab (SPARCL)

Table of contents

- Public version for SDSS DR16, DESI EDR
- DESI: healpix coadds for TARGETID>0
  (**no cuts** on COADD_FIBERSTATUS or ZWARN)
- Fast data discovery and access to 1D spectra (retrieval of up to 20k spectra/call)

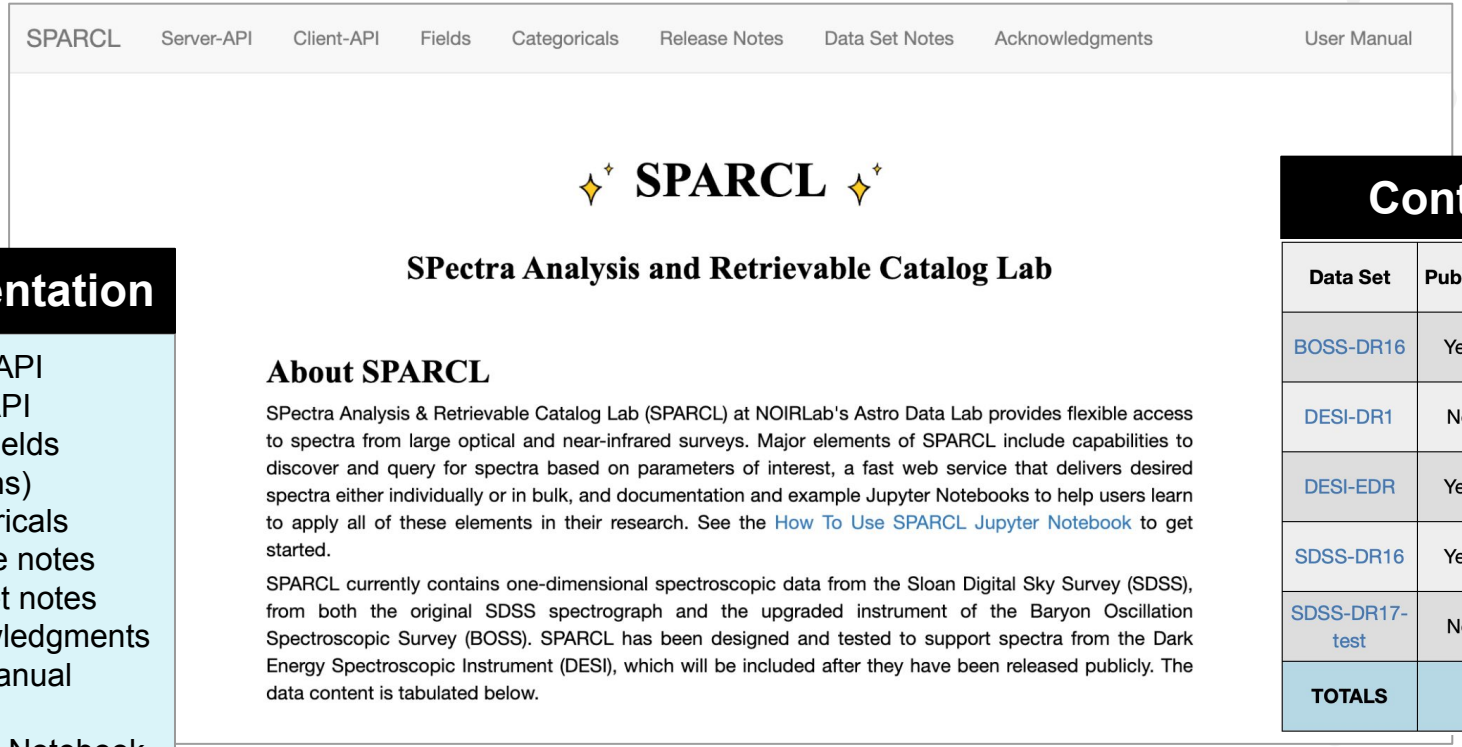
ADASS Proceedings arXiv:2401.05576

Full notebook available (Jacques et al.):
github.com/astro-datalab/notebooks-latest/blob/master/04_HowTos/SPARCL/How_to_use_SPARCL.ipynb

Tutorial notebooks for DESI spectra

- How_to_use_SPARCL.ipynb: Public version for SDSS DR16, DESI EDR available at Astro Data Lab (in user accounts and on GitHub)
- How_to_use_SPARCL_with_DESI_DR1.ipynb: DESI colleague version for DESI DR1 at Astro Data Lab (today!)
- How_to_use_SPARCL_at_NERSC.ipynb: DESI colleague version for DESI DR1 at NERSC (announced in Dec 2024)

Website



Documentation

- Server API
- Client API
- List of fields (columns)
- Categoricals
- Release notes
- Data set notes
- Acknowledgments
- User Manual

\+ How-To Notebook

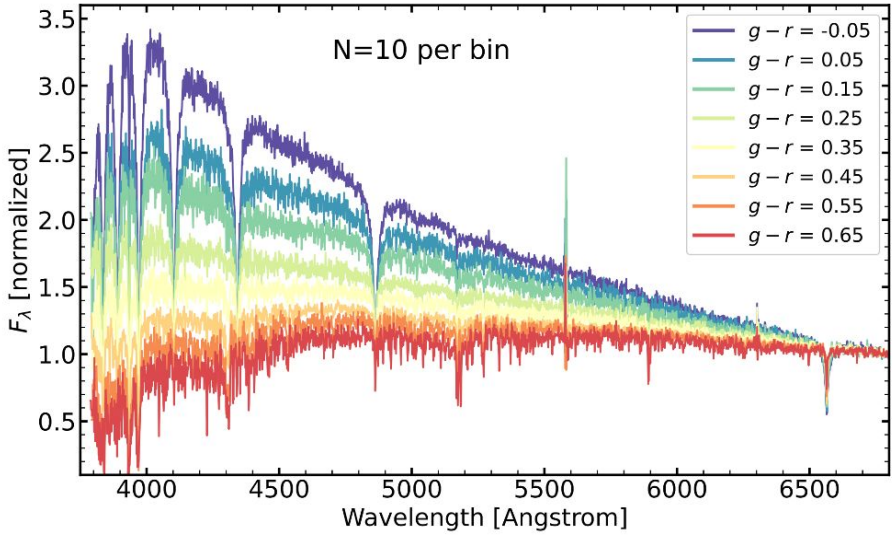
Contents

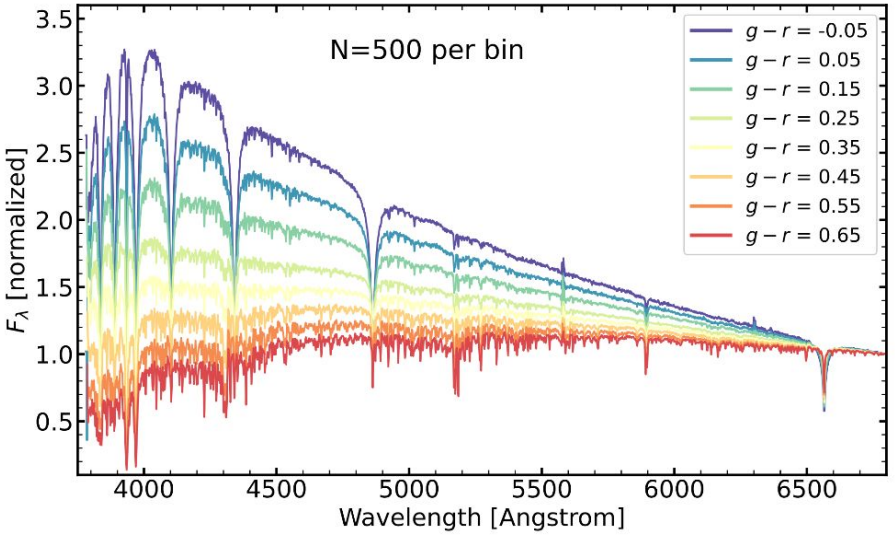
Data Set	Public?	# of Records Total
BOSS-DR16	Yes	3,946,000
DESI-DR1	No	23,060,727
DESI-EDR	Yes	2,044,588
SDSS-DR16	Yes	1,843,200
SDSS-DR17-test	No	1,843,200
TOTALS		**32,737,715**

astrosparcl.datalab.noirlab.edu

Example SPARCL use case

1) Using the Astro Data Lab, query “N” stars in eight bins of *g* – *r* color
2) Using SPARCL, retrieve and stack SDSS spectra in each color bin
Results: (left) N=10 takes 5sec for 80 spectra; (right) N=500 takes 90sec for 4000 spectra

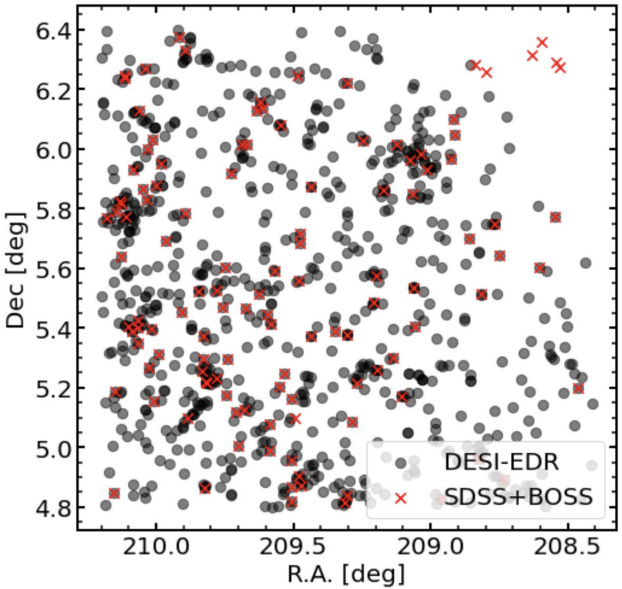




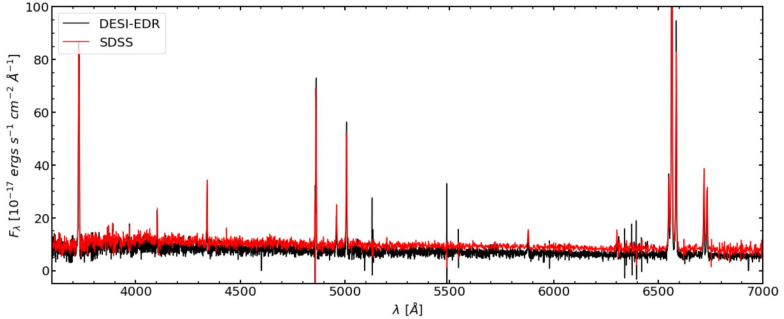
(Figure from Juneau et al. 2024; ADASS Proceedings arXiv:2401.05576)

DESI & SDSS Comparison Notebook (Pucha et al.)

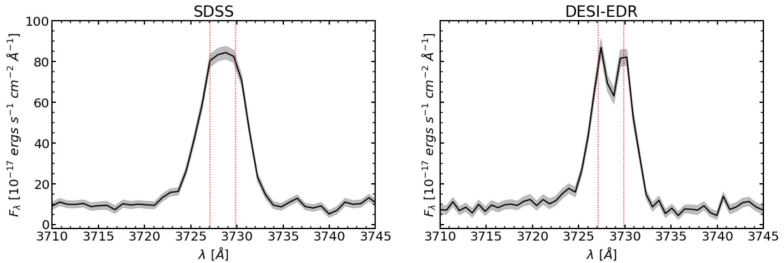
1) Use SPARCL to find SDSS and DESI-EDR spectra
2) Select example with high [OII] flux from SDSS
3) Compare spectra and [OII] doublet



From this figure, we clearly see that the DESI target density is much higher than SDSS.




```
fig, axs = plt.subplots(1,2,figsize = (20,6), sharex = True, sharey = True)

axs[0].plot(lam_sdss, flux_sdss, color = 'k', lw = 2.0)
axs[0].fill_between(lam_sdss, flux_sdss-err_sdss, flux_sdss+err_sdss, color = 'grey', alpha = 0.5)
axs[1].plot(lam_desi, flux_desi, color = 'k', lw = 2.0)
axs[1].fill_between(lam_desi, flux_desi-err_desi, flux_desi+err_desi, color = 'grey', alpha = 0.5)
axs[0].axvline(3727.092, color = 'r', ls = ':')
axs[0].axvline(3729.875, color = 'r', ls = ':')
axs[1].axvline(3727.092, color = 'r', ls = ':')
axs[1].axvline(3729.875, color = 'r', ls = ':')

axs[0].set(xlabel = '$\lambda~[\AA]$', ylabel = '$F_{\lambda}~[10^{-17}~ergs~s^{-1}~cm^{-2}~{\AA}^{-1}]$', title = 'SDSS')
axs[1].set(xlim = [3710,3745], ylim = [-2,100], title = 'DESI-EDR', xlabel = '$\lambda~[\AA]$', ylabel = '$F_{\lambda}~[10^{-17}~ergs~s

plt.show()
```



Resources

SPARCL

SPARCL website: https://astrosparcl.datalab.noirlab.edu

Contact: datalab-spectro@noirlab.edu

Astro Data Lab

Astro Data Lab website: https://datalab.noirlab.edu

Register account: https://datalab.noirlab.edu/account/register.html

User manual: https://datalab.noirlab.edu/docs/manual

Helpdesk: https://datalab.noirlab.edu/help

Notebook collection: https://github.com/astro-datalab/notebooks-latest

Extra slides

Notebooks: How to use SPARCL

How to use SPARCL

SPectra Analysis and Retrievable Catalog Lab (SPARCL)

Table of contents

- Public version for SDSS DR16, DESI EDR
- DESI: healpix coadds for TARGETID>0
  (**no cuts** on COADD_FIBERSTATUS or ZWARN)
- Fast data discovery and access to 1D spectra (retrieval of up to 20k spectra/call)

ADASS Proceedings arXiv:2401.05576

Full notebook available (Jacques et al.):
github.com/astro-datalab/notebooks-latest/blob/master/04_HowTos/SPARCL/How_to_use_SPARCL.ipynb

Notebook: How to use SPARCL

Discover data from SPARCL client

Data discovery I: using SPARCL's `client.find()` method

The first way you can discover your data is by using SPARCL's `client.find()` method, which allows you to find records in the SPARCL database based on certain parameters passed to the function. Only Core fields may be in the `outfields` and `constraints` parameters. The descriptions for all fields, including Core fields, is located here. The SPARCL Core fields constraint types are:

Field name	Constraint type	Example
id	List of values (but not intended for data discovery)	['00001658-460c-4da1-987d-e493d8c9b89b', '000017b6-56a2-4f87-8828-3a3409ba1083']
specid	List of values	[6988698046080241664, 6971782884823945216]
targetid	List of values	[1237679502171374316, 1237678619584692841]
data_release	List of allowed values from SPARCL Categoricals	['BOSS-DR16', 'SDSS-DR16']
datasetgroup	List of allowed values from SPARCL Categoricals	['SDSS_BOSS']
ra	Range of values (may not "wrap" around RA=0)	[44.53, 47.96]
dec	Range of values	[2.03, 7.76]
redshift	Range of values	[0.5, 0.9]
redshift_err	Range of values	[0.000225, 0.000516]
redshift_warning	List of values	[0, 3, 5]
spectype	List of allowed values from SPARCL Categoricals	['GALAXY', 'STAR']
instrument	List of allowed values from SPARCL Categoricals	['SDSS', 'BOSS']

Discover data from Data Lab database

Data discovery II: using Data Lab's `sparcl.main` table

The second way you can discover your data is by querying the `sparcl.main` table hosted at Data Lab. columns:

```
print(qc.schema('sparcl.main'))
```

```
Schema: sparcl
 Table: main

     Column Name   Description
     -----------   -----------
            elon   Ecliptic longitude
            elat   Ecliptic latitude
            glon   Galactic longitude
            glat   Galactic latitude
        redshift   Observed redshift, or radial velocity in units of redshift
              ra   Right Ascension in degrees
             dec   Declination in degrees
         wavemin   Minimum value of wavelength coverage
         wavemax   Maximum value of wavelength coverage
    redshift_err   Uncertainty on the observed redshift, or radial velocity i
                   n units of redshift
        targetid   Dataset-specific photometric target identifier
          specid   Dataset-specific spectrum identifier (may not be unique)
         exptime   Exposure time in seconds
       random_id   Random ID in the range 0 to 100 (decimal)
            htm9   HTM Level-9 index
         ring256   Healpix ID with nside 256, ring schema
        nest4096   Healpix ID with nside 4096, nested schema
redshift_warning   Bitmask with warning on redshift result
```

Notebook: How to use SPARCL

Convert to Astropy Spectrum1D object

Spectrum1D:

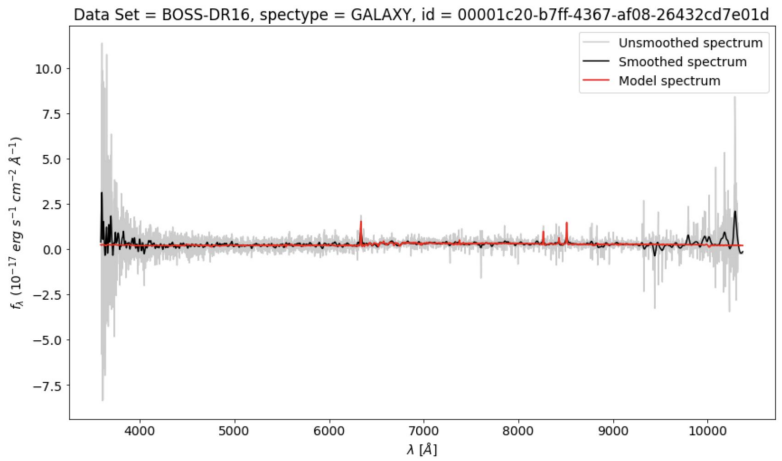
When applicable, the units for each field are documented on the Fields tab of the astrosparcl website

```
specs = [Spectrum1D(spectral_axis = r.wavelength*u.AA,
                    flux = np.array(r.flux)* 10**-17 * u.Unit('erg cm-2 s-1 AA-1'),
                    uncertainty = InverseVariance(np.array(r.ivar)),
                    redshift = r.redshift,
                    mask = r.mask)
        for r in results_I.records]
specs[0]
```

```
<Spectrum1D(flux=<Quantity [ 8.68927574e-17,  8.68837833e-17,  8.68748474e-17, ...,
           1.01688528e-17, -8.76587963e-17, -8.76609707e-17] erg / (Angstrom cm2 s)>, spectra
l_axis=<SpectralAxis
   (observer to target:
      radial_velocity=115835.56522734222 km / s
      redshift=0.503122508525848)
  [ 3562.87019405,  3563.69066968,  3564.51133426, ..., 10389.62804909,
   10392.0206248 , 10394.41375149] Angstrom>, uncertainty=InverseVariance([0.        , 0.
, 0.        , ..., 0.        ,
                0.19256701, 0.        ]))>
```

Plot spectrum + best-fit model

```
plot_spec(index=2, results=results_I)
```



Full notebook available (Jacques et al.):
github.com/astro-datalab/notebooks-latest/blob/master/04_HowTos/SPARCL/How_to_use_SPARCL.ipynb

DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

How to use SPARCL at NERSC

SPectra Analysis & Retrievable Catalog Lab

Tutorial Jupyter notebook demonstrates how to:

- **Install** SPARCL locally (sparclclient)
- **Authenticate** with your email/password
- **Find** objects/spectra based on zpix or other catalogs
  - Search from `iron` specprod database (*fast!*)
- Quickly **retrieve** selected spectra
  - 20 spectra in 1.5 sec
  - 1000 spectra in ~30 sec
- **Convert** to other formats (e.g., Spectrum1D)
- **Plot** some selected spectra
  - using the *mask* to plot valid data
  - overlay best-fit *model* (Redrock template)

SPARCL

stephanie.juneau@noirlab.edu
datalab-spectro@noirlab.edu

AURA

Notebook: How to use SPARCL at NERSC

How to use SPARCL at NERSC

SPectra Analysis and Retrievable Catalog Lab (SPARCL)

Alice Jacques (NOIRLab), Stéphanie Juneau (NOIRLab), Benjamin Weaver (NOIRLab), Steve Pothier (NOIRLab), Adam Bolton (SLAC) and the SPARCL team

Table of contents

NERSC

`/global/cfs/cdirs/desi/science/gqp/sparcl/`

`How_to_use_SPARCL_at_NERSC.ipynb`

GitHub (coming soon!)

https://github.com/desihub/tutorials/

→ database

- How_to_use_SPARCL_at_NERSC.ipynb
- README.md
- spectroscopic-production-database.ipynb

PART II: Exploring DESI DR1 efficiently with the **Astro Data Lab** science platform

Value-added services to maximize discovery

Stéphanie Juneau

On behalf of the NSF NOIRLab Astro Data Lab & SPARCL teams

(SPARCL: A. Jacques, P. Peterson, B. Weaver, S. McManus, R. Nikutta, A. Bolton, R. Pucha)

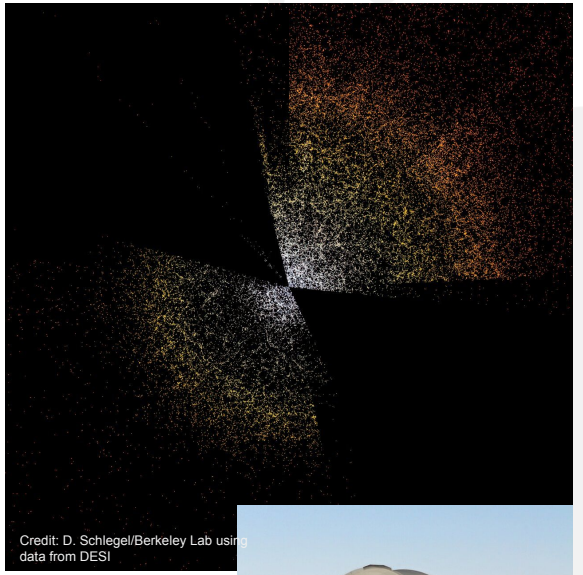
DESI (Dark Energy Spectroscopic Instrument)

General DESI website: www.desi.lbl.gov
DESI Data documentation website: data.desi.lbl.gov/doc/

- Primarily a DOE cosmology experiment (stage IV) but also a treasure trove of data for astronomers and students!

- Data Releases
  - DESI-EDR: Survey Validation; spectra of 1.7 million unique targets
  - DESI-DR1: Data up to 13 months (“Year 1”); spectra of 12.8 million unique targets (galaxies, quasars & stars)
  - Future data releases will keep growing (DR2 for Year 3 data)

- DESI at NOIRLab:
  - Data in the Astro Data Archive: astroarchive.noirlab.edu
  - Catalog Database in the Astro Data Lab: datalab.noirlab.edu
  - Spectral Database as part of SPARCL: astrosparcl.datalab.noirlab.edu


Credit: D. Schlegel/Berkeley Lab using data from DESI

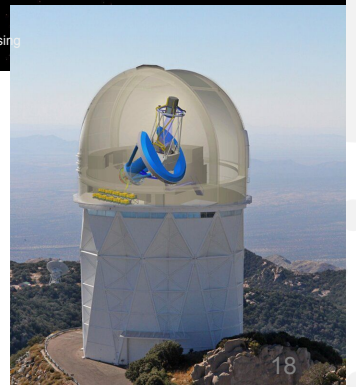
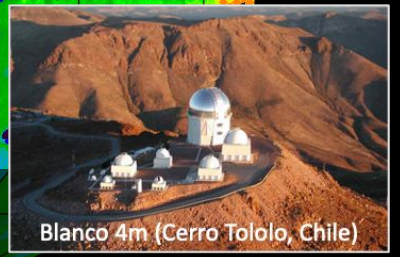


Image credit: R. Lafever, J. Moustakas/DESI Collaboration, P. Marenfeld/NOIRLab/AURA/NSF & E. Acosta/Vera C. Rubin Observatory/ NOIRLab/ NSF/ AURA

The Astro Data Lab

datalab.noirlab.edu



Mission

Efficient exploration and analysis of large astronomy datasets with an emphasis on NOIRLab wide-field 4-meter telescopes

The Astro Data Lab Science Platform

Pre-installed software/tools/tutorials co-located w/ the data

- Web services (website)
- JupyterLab (notebook) server

Low barrier of entry to powerful tools

Rich variety & volume of data: images (2.5 PB), catalogs (175 TB) (+databases with >219 B rows), spectra (40+ M) from ground-based (primary focus) and space-based observatories

Access to big data

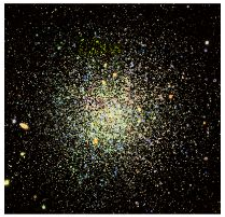
Astronomers/students can create a user account, log on, use our services for their *entire* analysis (directly from their browser or install a command-line package)

Open & inclusive
Community oriented

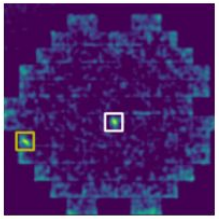
For DESI: support the overall success and first impression of the DESI survey data by facilitating a user-friendly access including for researchers without NERSC credentials

Discovery Ready Science Notebooks

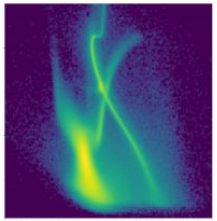
Astro Data Lab Notebook Gallery



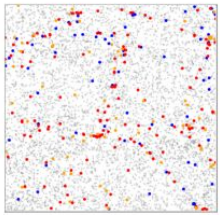
Getting Started with Data Lab
Learn the basics such as importing modules, sending a database query, and using the Simple Image Access (SIA) service to create image cutouts.



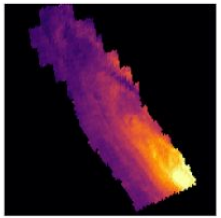
Dwarf Galaxies in the SMASH survey
Discover the ultrafaint Hydra II dwarf galaxy in the SMASH DECam survey based on spatial overdensities of blue stars with a detection algorithm.



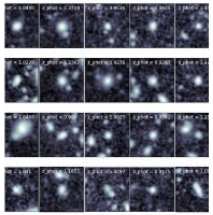
Exploring SMASH DR2
The 480 square degree SMASH DECam survey of the Magellanic Clouds and their periphery contains a wealth of objects, including this capture of the SMC with 47 Tuc in the foreground.



Large-Scale Structure of the Universe
Investigate cosmic filaments and clusters of galaxies, pan around an interactive sky viewer, combining spectroscopy and DESI pre-imaging.



Fun with PHAT
Visualize the 100 million+ stars in the Andromeda Galaxy captured by the Panchromatic Hubble Andromeda Treasury (PHAT).



Gallery of Cluster Galaxies
Use the Simple Image Access (SIA) service to retrieve images from the Gemini GOGREEN program.

Successes

- Training
- Education
- Research

Science cases

- Intro
- Galactic
- Extragalactic
- Time Domain

Datasets

- Images
- Catalogs
- Spectra

DESI at Astro Data Lab

DESI landing page

- Refers to DESI collaboration website(s), papers and acknowledgments
- Draft version available for comments [+screenshots]

DESI databases

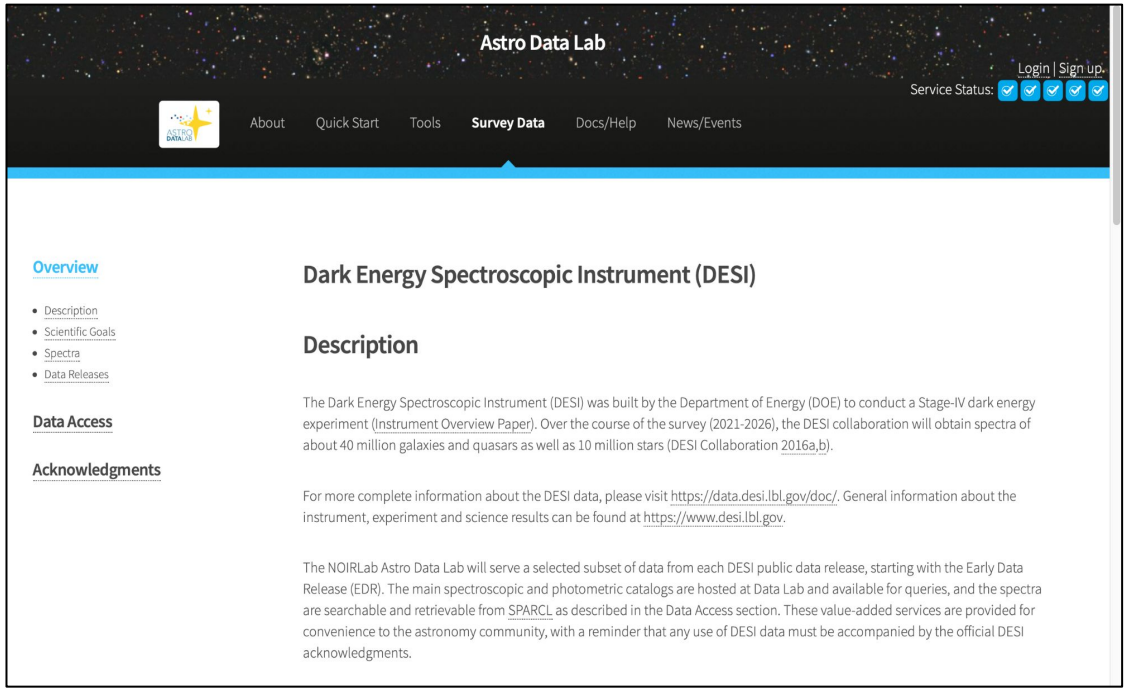
- **fuji** (NERSC) → **desi_edr**
- **iron** (NERSC) → **desi_dr1**

Tutorial notebooks

- DESI 23.1 (22.2) kernel installed to use with DR1 (EDR) notebooks
- *Getting started* notebook (EDR tutorial for now; DR1 at public release)
- Database tutorial notebook (adapted from tutorial by Ben Weaver; EDR version available)
- Science example:
  - find galaxies with both SDSS and DESI spectra; compare the spectra and [OII] doublet

DESI Landing Page

(not yet live) datalab.noirlab.edu/desi/desi.php

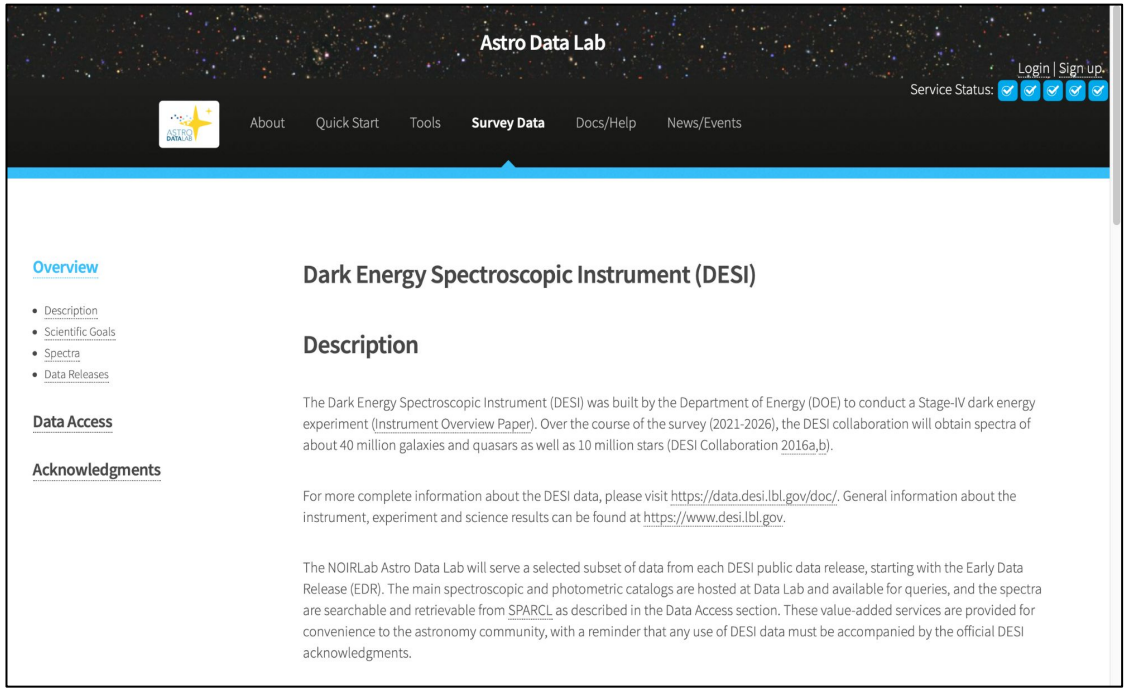


The Dark Energy Spectroscopic Instrument (DESI) was built by the Department of Energy (DOE) to conduct a Stage-IV dark energy experiment (Instrument Overview Paper). Over the course of the survey (2021-2026), the DESI collaboration will obtain spectra of about 40 million galaxies and quasars as well as 10 million stars (DESI Collaboration 2016a,b).

[...] These value-added services are provided for convenience to the astronomy community, with a reminder that any use of DESI data must be accompanied by the official DESI acknowledgments.

DESI Landing Page

(not yet live) datalab.noirlab.edu/desi/desi.php



Scientific Goals

The DESI collaboration seeks to understand the nature of dark energy and its evolution. In addition, the DESI data enable a suite of other scientific discoveries and research thanks to the legacy value of the immense spectroscopic dataset. The DESI Collaboration maintains an ADS library of DESI papers.

DESI Spectra

The DESI collaboration produces one-dimensional (1-D) spectra that have been sky-subtracted, flux-calibrated, and modeled with a best-fitting template to identify the spectral type (spectype) and measure the redshift (z) with the Redrock pipeline (Bailey et al., in preparation). There are a few categories of spectra at different stages of data combination. Namely:

- Per exposure (spectra from a single exposure at the telescope)
- Per night (coadded spectra obtained on multiple exposures during the same night)
- Per tile (coadded spectra obtained for the same tile but possibly from multiple nights)
- Per healpix (coadded spectra per position on the sky grouped by HEALPixels; which can be from multiple nights and/or multiple tiles)

Furthermore, the spectra are treated separately per SURVEY:

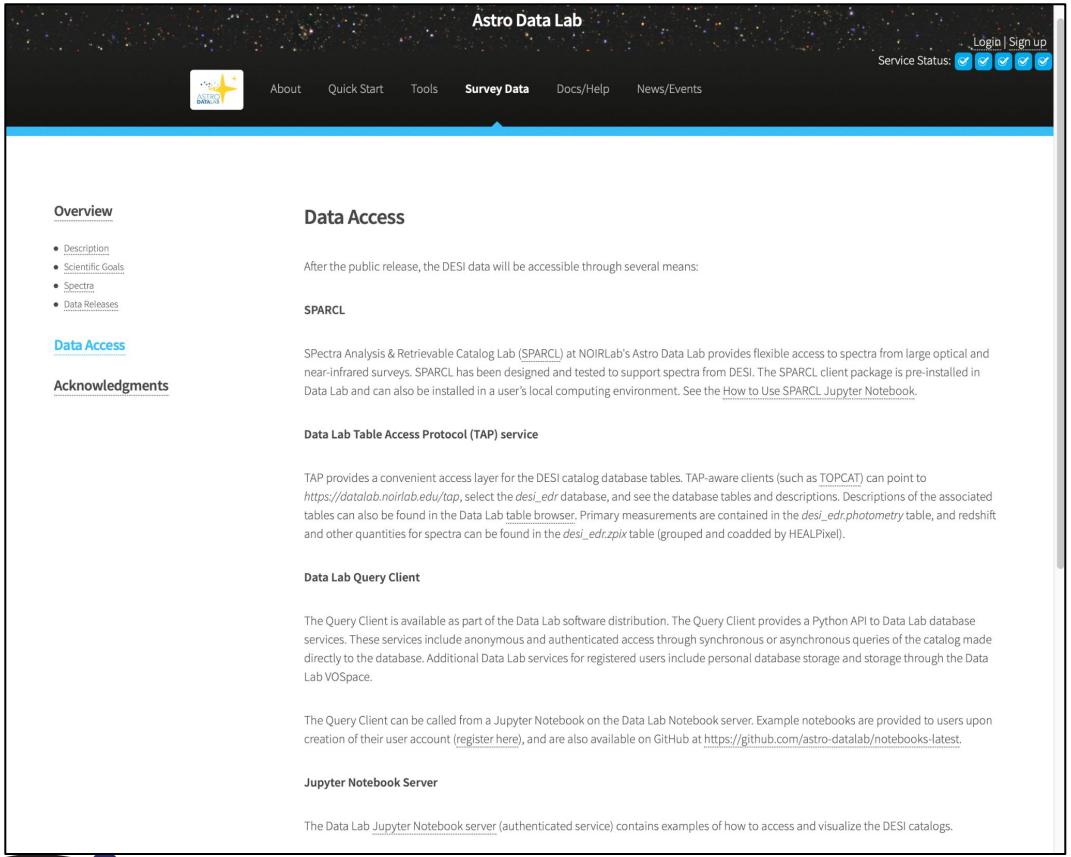
- cmx = commissioning
- sv1 = first phase of Survey Validation (SV)
- sv2 = operations testing
- sv3 = one-percent survey (covering 1% of area; final phase of SV)
- main = main DESI survey (starting with DR1)

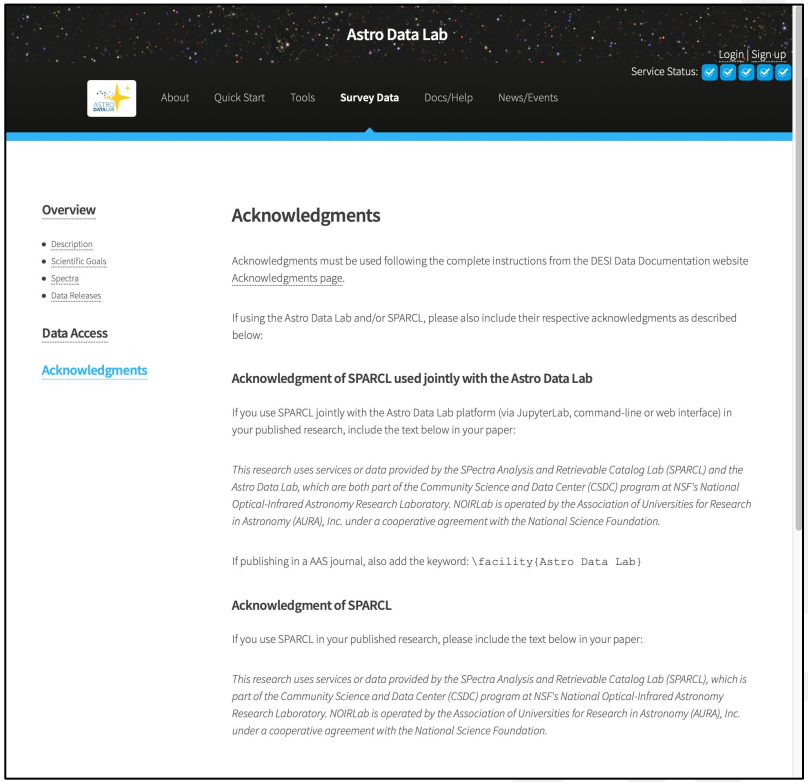
And per PROGRAM:

- dark = dark time for the main targets: Luminous Red Galaxies (LRG), Emission-Line Galaxies (ELG), and Quasi-Stellar Objects / Quasars (QSO)
- bright = bright time for the Bright Galaxy Survey (BGS) and the Milky Way Survey (MWS)

The tables listed below include a redshift catalog per healpix (**zpix**; *recommended for most analyses*) and a redshift catalog per tile (**ztile**; can be useful for instrument-dependent analyses). The SPARCL database only includes DESI spectra that have been coadded per healpix (HEALPixel).

DESI Landing Page





DESI Database at Astro Data Lab

desi_edr (fuji at NERSC)

Table Name	Description
exposure	Summary quantities for every DESI exposure
fiberassign	Quantities obtained when a DESI target is assigned to a fiber
frame	Summary quantities for each petal of the DESI instrument in a given exposure; in normal operation there are ten frames for every exposure
photometry	Photometric quantities from LS DR9 for every TARGETID
potential	For a given tile, this table lists all targets that could have received a fiber assignment
target	The quantities obtained when photometric objects are analyzed in the target selection process
tile	Summary quantities for every DESI tile (pointing on the sky)
zpix	Redshift and other quantities for spectra that have been grouped and coadded by HEALPixel
ztile	Redshift and other quantities for spectra that have been grouped and coadded by tile

+ Selected VACs (e.g., fastspecfit)

DESI-EDR Database

The `desi_edr` database schema at DataLab consists of several tables associated with the targets. The different tables that are available are:

```
print(qc.schema('desi_edr'))
```

```
Schema: desi_edr

     Table Name   Description
     ----------   -----------
```

We will focus on three different tables in this notebook:

1. `zpix` : This table is the main redshift catalog from the DESI survey. It contains the redshift information from the coadded-spectra, based on the healpix of the objects.
2. `target` : This table contains targeting information of the different sources.
3. `fiberassign` : This table contains the fiber assignment information of the targets.

```
query = ("SELECT * FROM desi_edr.zpix LIMIT 5")
t = qc.query(sql = query, fmt = 'table')

print (t.columns)
```

```
<TableColumns names=('id','targetid','survey','program','spgrp','spgrpval','healpix','z','zerr','zwarn','chi2','coeff_0','coeff_1','coeff_2','coeff_3','coeff_4','coeff_5','coeff_6','coeff_7','coeff_8','coeff_9','npixels','spectype','subtype','ncoeff','deltachi2','coadd_fiberstatus','coadd_numexp','coadd_exptime','coadd_numnight','coadd_numtile','mean_delta_x','rms_delta_x','mean_delta_y','rms_delta_y','mean_fiber_ra','std_fiber_ra','mean_fiber_dec','std_fiber_dec','mean_psf_to_fiber_specflux','tsnr2_gpbdark_b','tsnr2_elg_b','tsnr2_gpbbright_b','tsnr2_lya_b','tsnr2_bgs_b','tsnr2_gpbbackup_b','tsnr2_qso_b','tsnr2_lrg_b','tsnr2_gpbdark_r','tsnr2_elg_r','tsnr2_gpbbright_r','tsnr2_lya_r','tsnr2_bgs_r','tsnr2_gpbbackup_r','tsnr2_qso_r','tsnr2_lrg_r','tsnr2_gpbdark_z','tsnr2_elg_z','tsnr2_gpbbright_z','tsnr2_lya_z','tsnr2_bgs_z','tsnr2_gpbbackup_z','tsnr2_qso_z','tsnr2_lrg_z','tsnr2_gpbdark','tsnr2_elg','tsnr2_gpbbright','tsnr2_lya','tsnr2_bgs','tsnr2_gpbbackup','tsnr2_qso','tsnr2_lrg','sv_nspec','sv_primary','main_nspec','main_primary','zcat_nspec','zcat_primary')>
```

```
query = ("SELECT * FROM desi_edr.target LIMIT 5")
t = qc.query(sql = query, fmt = 'table')

print (t.columns)
```

```
<TableColumns names=('id','targetid','photsys','subpriority','obsconditions','priority_init','numobs_init','hpxpixel','cmx_target','desi_target','bgs_target','mws_target','sv1_desi_target','sv1_bgs_target','sv1_mws_target','sv2_desi_target','sv2_bgs_target','sv2_mws_target','sv3_desi_target','sv3_bgs_target','sv3_mws_target','scnd_target','sv1_scnd_target','sv2_scnd_target','sv3_scnd_target','survey','program','tileid')>
```

```
query = ("SELECT * FROM desi_edr.fiberassign LIMIT 5")
t = qc.query(sql = query, fmt = 'table')

print (t.columns)
```

```
<TableColumns names=('id','tileid','targetid','petal_loc','device_loc','location','fiber','fiberstatus','target_ra','target_dec','lambda_ref','fa_target','fa_type','fiberassign_x','fiberassign_y','priority','plate_ra','plate_dec')>
```

DESI Notebooks at Astro Data Lab

DESI EDR Tutorial Notebook (similar to desihub/tutorials/getting_started/intro_to_DESI_EDR_files.ipynb)

Accessing DESI Targets

The DESI spectra are uniquely identified by three quantities:

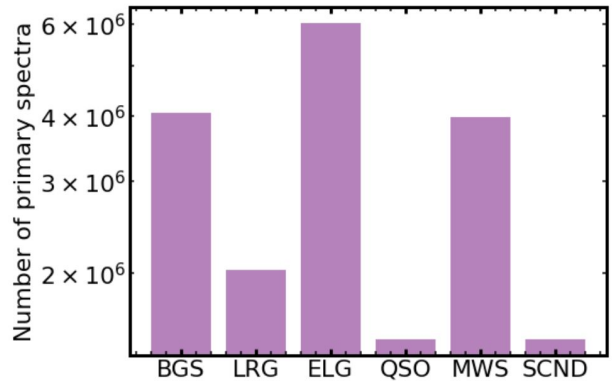
- **targetid**: unique identifier for a given target
- **survey**: SURVEY that the target was observed in. This can be cmx, sv1, sv2, or sv3.
- **program**: FIBER ASSIGNMENT PROGRAM. This is the planned observing conditions for the target. It can be *dark* or *bright* or *backup*. In case of cmx and sv1, there is *other* as well.

The different spectra observations of individual targets are coadded within each survey and program. Therefore, some targets may have multiple coadded spectra and each one is associated with a given targetid, survey, and program. The "best" spectrum for a given object is given by the `ZCAT_PRIMARY` column. More information about the data model of the redshift catalogs is available here.

```
query = """
SELECT zp.targetid, zp.survey, zp.program, zp.healpix,
       zp.z, zp.zwarn, zp.coadd_fiberstatus, zp.spectype, zp.zcat_nspec, zp.zcat_primary,
       tgt.desi_target, tgt.sv1_desi_target, tgt.sv2_desi_target, tgt.sv3_desi_target,
       fa.target_ra, fa.target_dec
FROM desi_edr.zpix AS zp
JOIN desi_edr.target AS tgt
ON (zp.targetid = tgt.targetid) AND (zp.survey = tgt.survey) AND (zp.program = tgt.program)
JOIN desi_edr.fiberassign as fa
ON (zp.targetid = fa.targetid)
WHERE (zp.zcat_primary = 't')
"""

## targetid, survey, program -- unique identifiers for a given spectrum
## healpix -- healpix number for the target
## z -- spectroscopic redshift of the target
## zwarn -- encoded information regarding the redshift (zwarn = 0 is good)
## coadd_fiberstatus -- encoded information regarding the fiber that is assigned to the target (coadd_fiberstatus = 0 is good)
## spectype -- Spectral type of the target: STAR | GALAXY | QSO
## zcat_nspec -- Number of coadded spectra that are available for a given target
## zcat_primary -- Whether or not a given coadded spectra is the primary spectra. ZCAT_PRIMARY = True for the "best" spectrum.
## desi_target -- encodeds main survey's desi targeting information - explained in detail below
## sv1_desi_target -- encodes sv1 desi targeting information
## sv2_desi_target -- encodes sv2 desi targeting information
## sv3_desi_target -- encodes sv3 desi targeting information
## target_ra, target_dec -- R.A. and Dec. of the target in degrees
```

```
# Number of sources of each target type
n_bgs = len(zpix_cat[is_bgs])
n_lrg = len(zpix_cat[is_lrg])
n_elg = len(zpix_cat[is_elg])
n_qso = len(zpix_cat[is_qso])
n_mws = len(zpix_cat[is_mws])
n_scnd = len(zpix_cat[is_scnd])
```

```
# Let us look at the numbers visually -

plt.figure(figsize = (8,6))

targets = ['BGS', 'LRG', 'ELG', 'QSO', 'MWS', 'SCND']
numbers = [n_bgs, n_lrg, n_elg, n_qso, n_mws, n_scnd]

plt.bar(targets, numbers, color = 'purple', alpha = 0.5)
plt.ylabel('Number of primary spectra')
plt.yscale('log')
```



DESI EDR Tutorial Notebook (cont'd)

Accessing and Plotting the Spectra of a Given Object

Finally, we show how to access all the available coadded spectra of a given object. We also show how to select the "best" spectrum.
We use the spectral access software, **SPARCL**, which is a fast spectral access service at DataLab.
Detailed ways to using SPARCL is available in this notebook. (**Link to the notebook**)

```
## Select STAR with nspec > 3
jj = (zpix_cat['zcat_nspec'] > 3)&(zpix_cat['spectype'] == 'STAR')
tsel = zpix_cat[jj]
```

```
len(tsel)
```

```
10056
```

```
## Randomly select an object

ii = 50
targetid = int(tsel['targetid'].data[ii])   ## SPARCL accepts only python integers
```

```
## Retrieve Spectra

inc = ['id', 'specid', 'redshift', 'flux', 'wavelength', 'spectype', 'specprimary', 'survey', 'program']
res = client.retrieve_by_specid(specid_list = [targetid],
                                include = inc,
                               dataset_list = ['DESI-EDR'])
```

```
## Checking that all the different spectra are retrieved

print ('Number of coadded spectra: ', tsel['zcat_nspec'].data[ii])
print ('Number of retrieved records: ', len(res.records))
```

```
Number of coadded spectra:  4
Number of retrieved records:  3
```

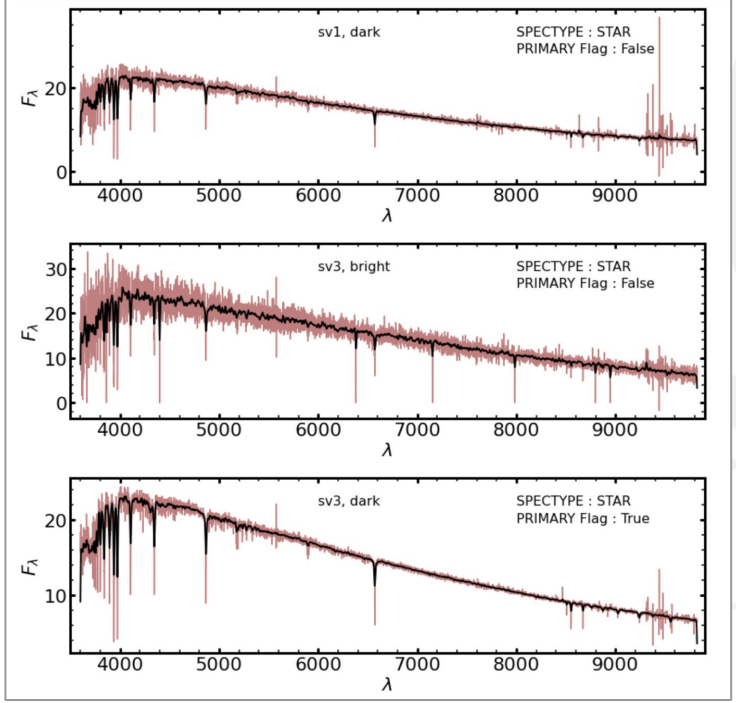
```
res
```

```
Retrieved Results: 3 records
```

```
records = res.records

## Select the primary spectrum
spec_primary = np.array([records[jj].specprimary for jj in range(len(records))])

primary_ii = np.where(spec_primary == True)[0][0]
```

```
lam_primary = records[primary_ii].wavelength
flam_primary = records[primary_ii].flux
```

```
# Number of spectra
n = len(records)

fig, ax = plt.subplots(n, 1, figsize = (12,(4*n)))

for jj in range(n):
    survey = records[jj].survey
    program = records[jj].program
    spectype = records[jj].spectype
    primary_flag = records[jj].specprimary

    lam = records[jj].wavelength
    flam = records[jj].flux

    # Plot the combined spectrum in maroon
    ax[jj].plot(lam, flam, color = 'maroon', alpha = 0.5)
    # Over-plotting smoothed spectra
    ax[jj].plot(lam, convolve(flam, Gaussian1DKernel(5)), color = 'k', lw = 2.0)
    ax[jj].set(xlim = [3500, 9900], xlabel = '$\lambda$', ylabel = '$F_{\lambda}$')
    trans = ax[jj].get_xaxis_transform()
    ax[jj].annotate(f'{survey}, {program}', xy = (6000, 0.85), xycoords = trans, fontsize = 16)
    ax[jj].annotate(f'SPECTYPE : {spectype}', xy = (8000, 0.85), xycoords = trans, fontsize = 16)
    ax[jj].annotate(f'PRIMARY Flag : {primary_flag}', xy = (8000, 0.75), xycoords = trans, fontsize = 16)

plt.tight_layout()
```



DESI EDR Tutorial Notebook (cont'd)

References

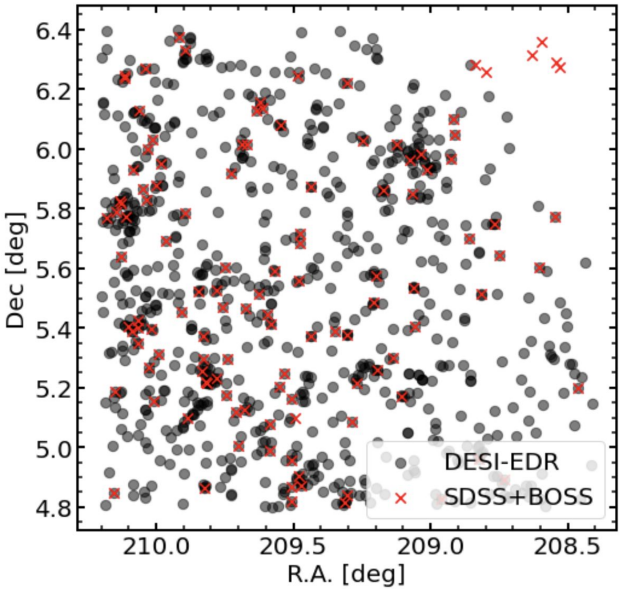
- Information about EDR: https://data.desi.lbl.gov/doc
- DESI Data License and Acknowledgments: https://data.desi.lbl.gov/doc/acknowledgements/
- Datamodel information: https://desidatamodel.readthedocs.io/en/latest/
  - tiles-fuji.fits
  - exposures-fuji.fits
  - Redshift catalogs
  - zall-pix-fuji.fits
  - zall-tilecumulative-fuji.fits
  - zpix-SURVEY-PROGRAM.fits
  - ztile-SURVEY-PROGRAM-GROUPTYPE.fits
  - coadd-SURVEY-PROGRAM-PIXNUM.fits
  - redshift-SURVEY-PROGRAM-PIXNUM.fits
- Information about DESI bitmasks: https://desidatamodel.readthedocs.io/en/latest/bitmasks.html

Acknowledgments

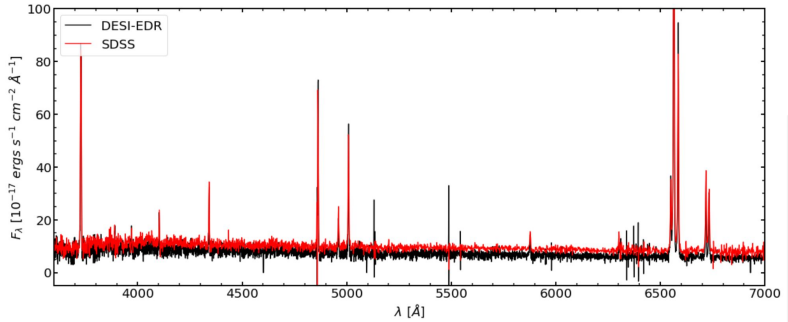
- DESI Data License and Acknowledgments: https://data.desi.lbl.gov/doc/acknowledgements/
- Astro Data Lab Acknowledgments: https://datalab.noirlab.edu/acknowledgements.php
- SPARCL Acknowledgments: https://astrosparcl.datalab.noirlab.edu/sparc/acknowledgments (TBC)

DESI & SDSS Comparison Notebook (Pucha et al.)

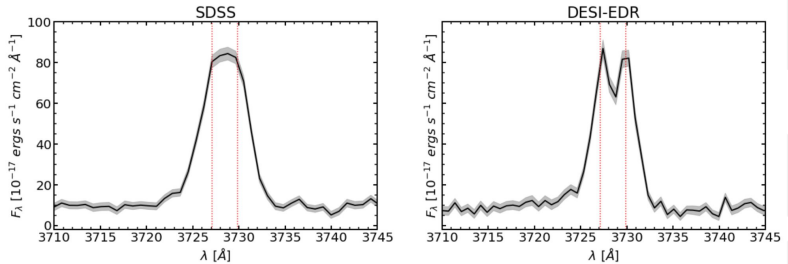
1) Use SPARCL to find SDSS and DESI-EDR spectra
2) Select example with high [OII] flux from SDSS
3) Compare spectra and [OII] doublet



From this figure, we clearly see that the DESI target density is much higher than SDSS.




```
fig, axs = plt.subplots(1,2,figsize = (20,6), sharex = True, sharey = True)

axs[0].plot(lam_sdss, flux_sdss, color = 'k', lw = 2.0)
axs[0].fill_between(lam_sdss, flux_sdss-err_sdss, flux_sdss+err_sdss, color = 'grey', alpha = 0.5)
axs[1].plot(lam_desi, flux_desi, color = 'k', lw = 2.0)
axs[1].fill_between(lam_desi, flux_desi-err_desi, flux_desi+err_desi, color = 'grey', alpha = 0.5)
axs[0].axvline(3727.092, color = 'r', ls = ':')
axs[0].axvline(3729.875, color = 'r', ls = ':')
axs[1].axvline(3727.092, color = 'r', ls = ':')
axs[1].axvline(3729.875, color = 'r', ls = ':')

axs[0].set(xlabel = '$\lambda~[\AA]$', ylabel = '$F_{\lambda}~[10^{-17}~ergs~s^{-1}~cm^{-2}~{\AA}^{-1}]$', title = 'SDSS')
axs[1].set(xlim = [3710,3745], ylim = [-2,100], title = 'DESI-EDR', xlabel = '$\lambda~[\AA]$', ylabel = '$F_{\lambda}~[10^{-17}~ergs~s

plt.show()
```



Astro Data Lab and SPARCL References

Astro Data Lab: https://datalab.noirlab.edu

Jupyter notebook server: https://datalab.noirlab.edu/devbooks/

All notebooks: https://github.com/astro-datalab/notebooks-latest/

All datasets: https://datalab.noirlab.edu/survey.php

User helpdesk / FAQ: https://datalab.noirlab.edu/help/

SPARCL: https://astrosparcl.datalab.noirlab.edu

Fitzpatrick, M. J., et al. 2014, *The NOAO Data Laboratory: a conceptual overview*, 2014SPIE.9149E..1TF

Nikutta, R., et al. 2020, *Data Lab – A community science platform*, 2020A&C....3300411N

Juneau, S., et al. 2021, *Jupyter-Enabled Astrophysical Analysis Using Data-Proximate Computing Platforms*, 2021CSE....23b..15J