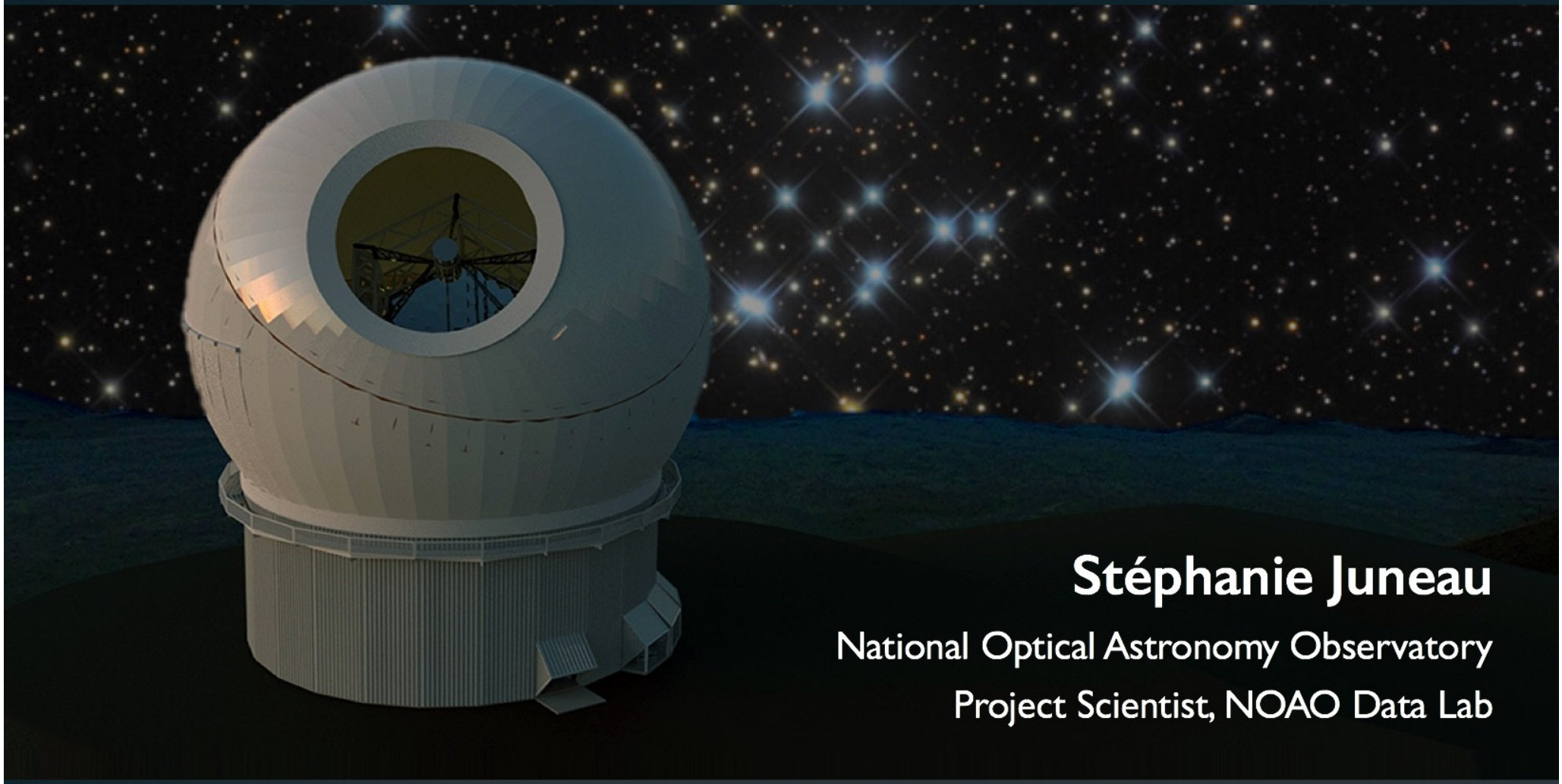
Wide field spectroscopy
in the future data ecosystem



Stéphanie Juneau

National Optical Astronomy Observatory

Project Scientist, NOAO Data Lab

Maunakea Spectroscopic Explorer

National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center

AURA

Wide field spectroscopy
in the future data ecosystem

We're not ready…

… but we can be!

Maunakea Spectroscopic Explorer

National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center

AURA

Astrophysical Data Ecosystem

- Astrophysical data have been growing in volume *and* complexity

- These trends will continue with current & upcoming missions and instruments, including MSE

→ Growing requirements in terms of hardware/software
→ Need to bring the analysis close to the data
→ Increase in Archive-driven and Data-driven science

- New mode(s) of doing science!

Need for Science Platforms

- Go beyond archiving data & making them available for download
- **Deliver analysis tools**: default toolkit (+user contributed?)
- Create collaborative work spaces
- Help solve diversity/accessibility problems

How?

- Let's share workload & wisdom
  - Interoperability between projects/archives/centers ?
  - Coordinated (if not centralized) effort !
  - Effort needs to be planned / intentional / funded (i.e., not an after-thought or assuming "software comes for free")

Need for Science Platforms

Several Science Platforms ongoing at data centers (e.g., CADC, IRSA, MAST, NRAO, NOAO Data Lab, CDS, …)

Status for spectroscopy:

Need for Science Platforms

Status for spectroscopy:

The Good:

- Some successful project examples (e.g., SDSS)
- Ongoing effort for spectroscopy software/tools (e.g., Astropy – Erik Tollerud's talk)

The Bad:

- Existing tools may not "scale" to the much larger surveys (DESI, MSE, Euclid, etc.)
- Still lacking a complete suite of open-source spectral reduction & analysis (Python) package

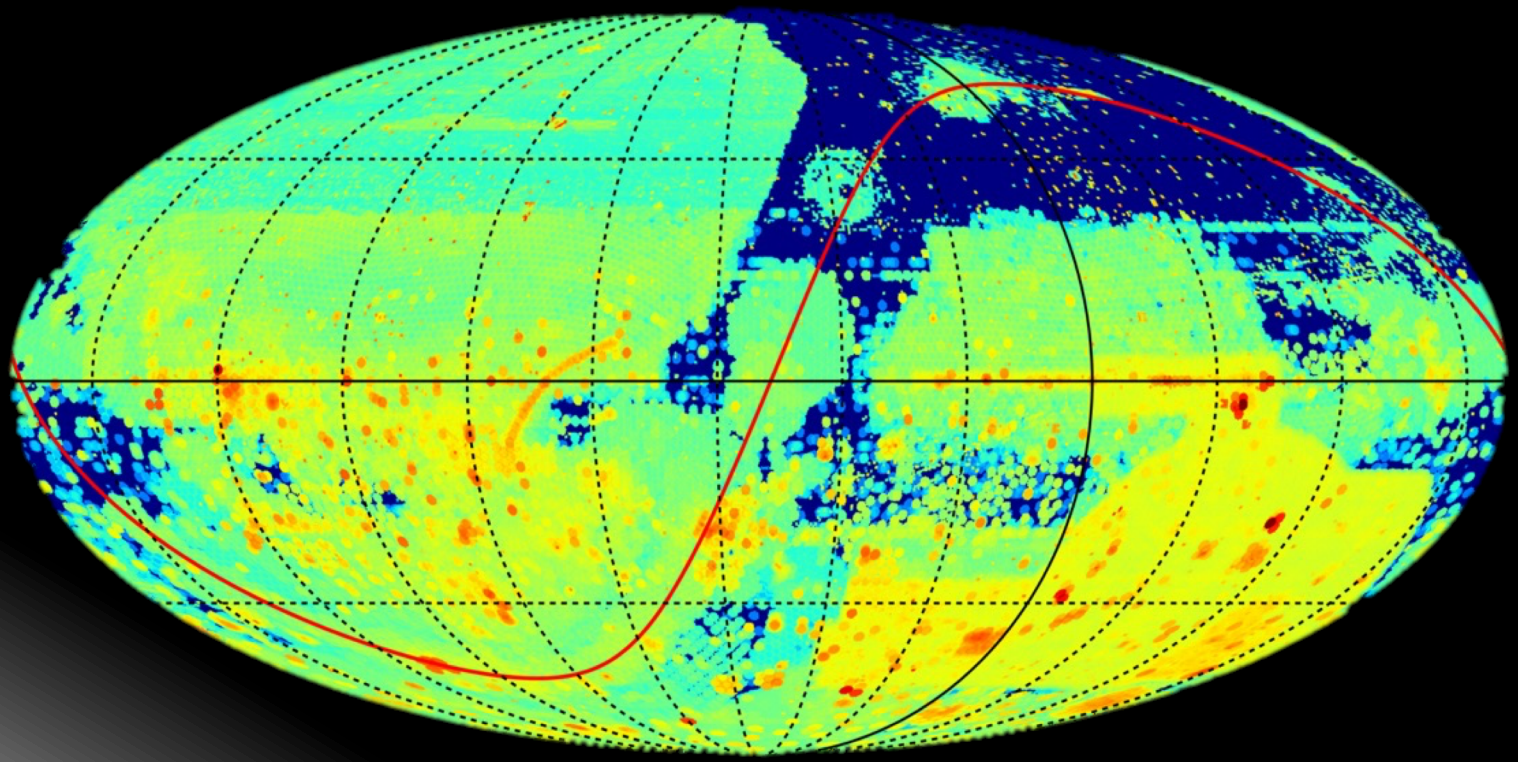
Need for Science Platforms

Status for spectroscopy:

The Wonderful:

- Many (all?) pieces of the puzzle exist
- We can create an (ambitious) plan to put them together:
  - Catalog/Database access
  - Contextual Viewer
  - Spectral Viewer
  - Analysis of reduced spectra
  - (Stacking reduced spectra – low S/N regime)
  - (Forward modeling and/or Machine Learning)
  - (Re-processing selected subset of spectra?)

The NOAO Data Lab

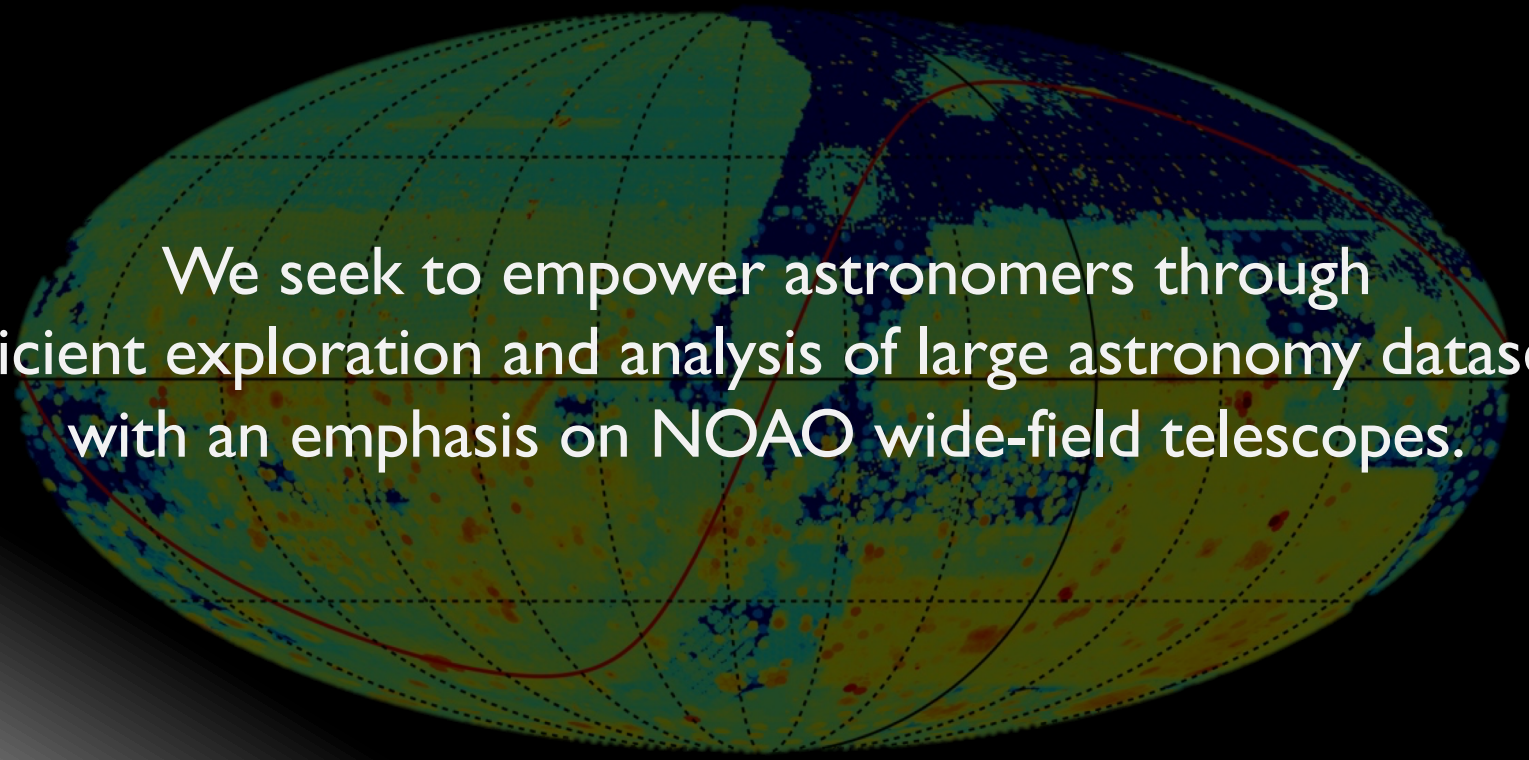


NOAO DATALAB

datalab.noao.edu

National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center
AURA
NSF

The NOAO Data Lab



We seek to empower astronomers through
efficient exploration and analysis of large astronomy datasets
with an emphasis on NOAO wide-field telescopes.

NOAO DATALAB

`datalab.noao.edu`

National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center
AURA
NSF

Datasets at the NOAO Data Lab

datalab.noao.edu

NOAO Facilities Featured Surveys:

DESI imaging Legacy Survey (LS): >1 billion objects in DR6+7

SMASH: ~100 million objects in DR1

DES: ~400 million objects in DR1

DECaPS: ~2 billion objects

NOAO All-Sky Source Catalog (NSC): ~2.9 billion objects

+ NSC “single-epoch”: ~30 billion measurements

Additional Surveys:

select tables from SDSS/BOSS DR13 & DR14, GAIA DR1 & DR2, DES SVA1, the Allen NEO catalog (DAD), and USNO-A2/B, *skinny* Pan-STARRS DR1, etc.

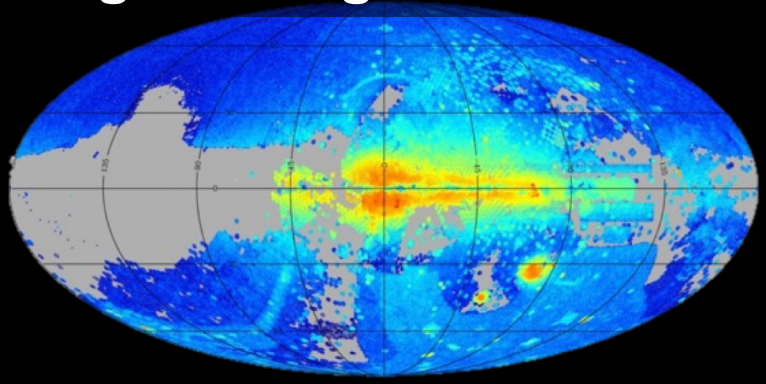
Summary of Current Functions

Function	Method
Sky exploration	Image discovery tool Catalog overlay tool
Authentication	Web interface datalab command Python authClient, DL interface
Catalog query	Web interface datalab command line (CLI) Python queryClient, DL interface TOPCAT
Image query	Simple Image Access (SIA) service
Query result storage	myDB Virtual storage space
File transfer	datalab command and Virtual storage space
Analysis	Jupyter notebook server

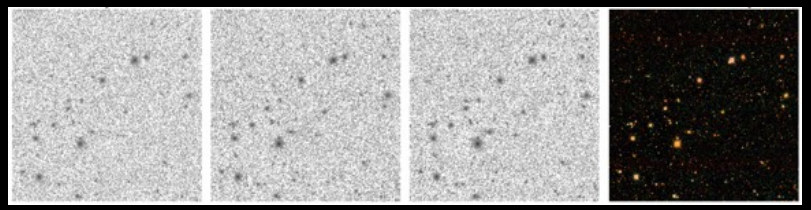
The NOAO Data Lab

datalab.noao.edu

Large Catalogs – TB-scale databases



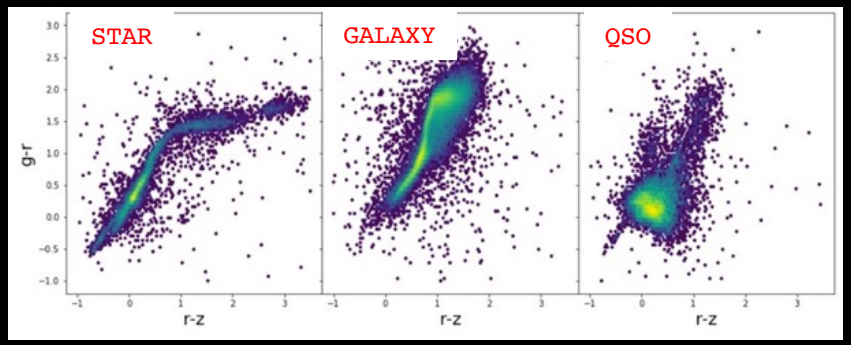
Pixel Data – images in NOAO Science Archive



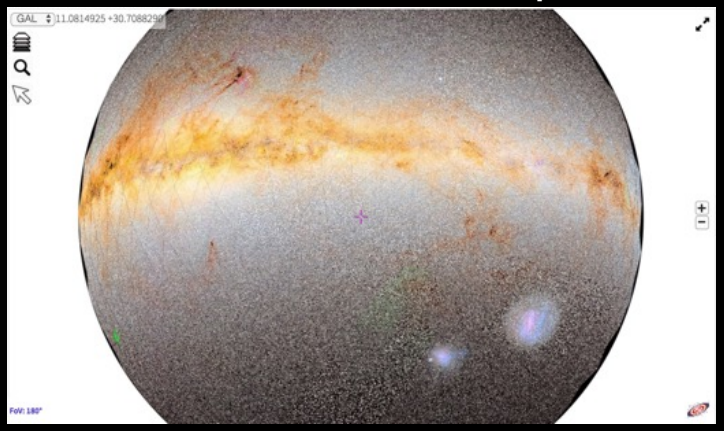
NOAO DATALAB

Virtual Storage – 1 TB per user

Compute Processing – workflows close to data



Visualization – data exploration



National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center
AURA

Legacy Surveys (LS) + SDSS/BOSS

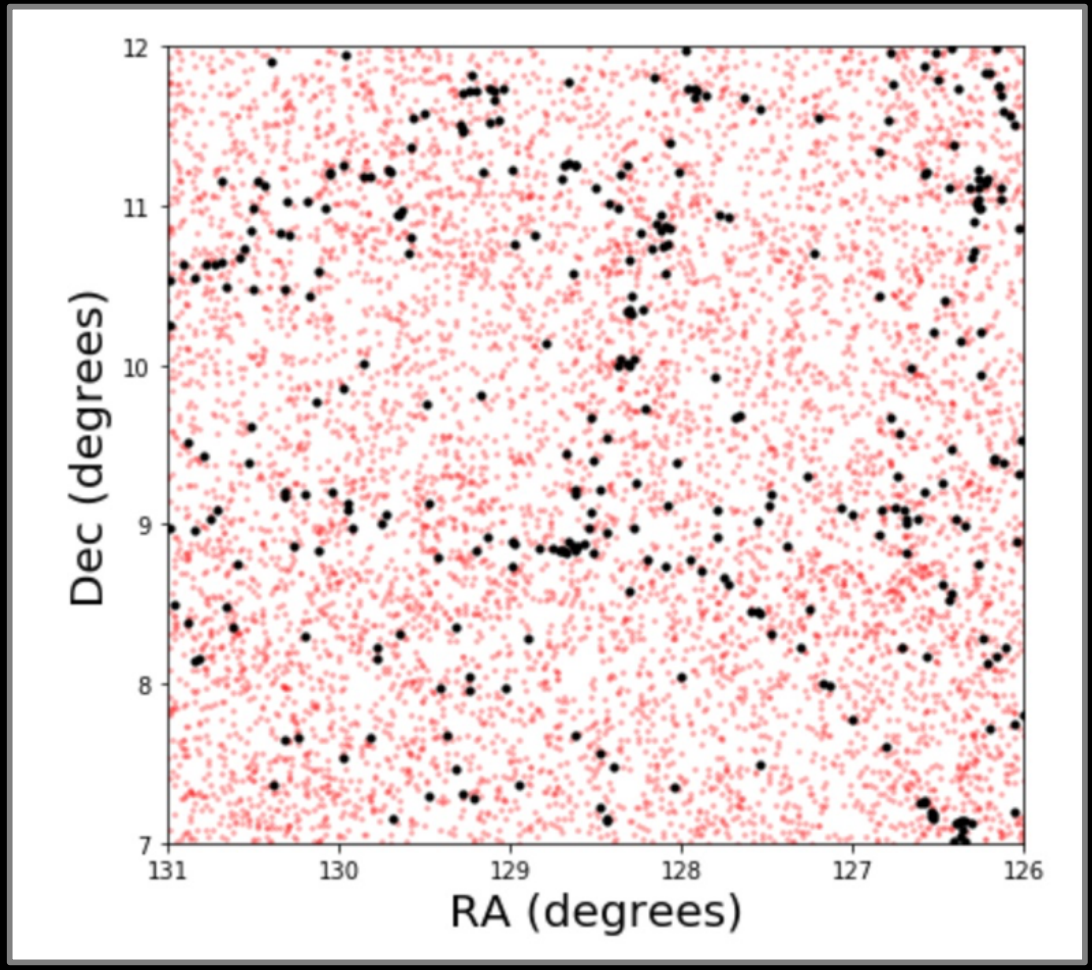
Example Science case:

Large-Scale Structure

Combine SDSS/BOSS spectroscopy with LS photometry to inspect galaxies along large-scale structures

See full notebook:

http://datalab.noao.edu/notebooks/web/ScienceExamples/LargeScaleStructure/LargeScaleStructureSdssLs_20180108.html



Red = all redshifts

Black = thin z slice (~similar distance)

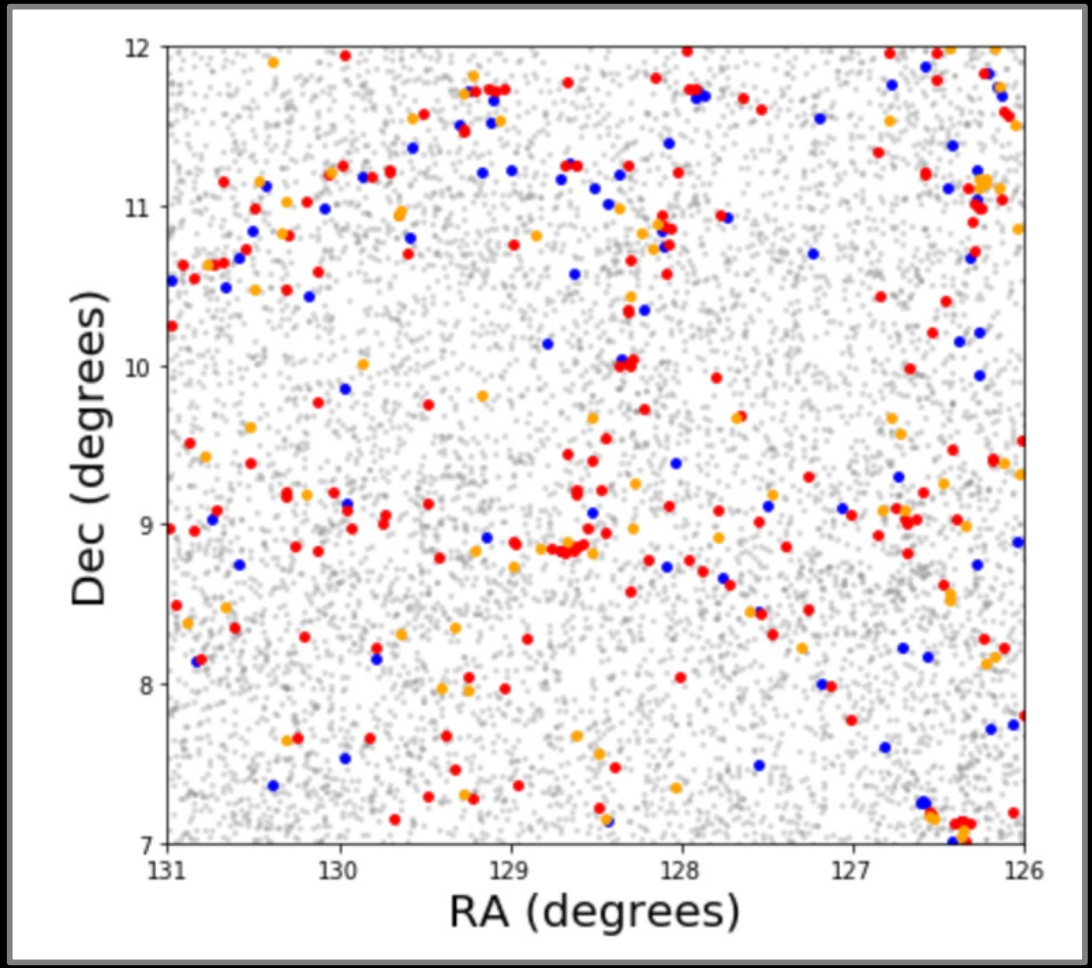
Legacy Surveys (LS) + SDSS/BOSS

Example Science case:

Large-Scale Structure

Combine SDSS/BOSS spectroscopy with LS photometry to inspect galaxies along large-scale structures

→ We can visually spot the *morphology-density* relation



morphology

- Red = de Vaucouleurs (elliptical)
- Orange = Bulge+Disk
- Blue = Exponential disk

DESI Dark Energy Spectroscopic Instrument

- 14,000 square degrees
- 35 million spectra of galaxies and quasars!
- 10 million spectra of stars
- Commissioning in 2019 (survey 2020-2025)

Object Class	Number of Spectra	Redshift Range
bright galaxies, $r < 19.5$	10 million	$0 < z < 0.4$
luminous red galaxies (LRGs)	4.2 million	$0.4 < z < 1.0$
emission line galaxies (ELGs)	18 million	$0.6 < z < 1.6$
quasars (QSOs)	2.4 million	$0.5 < z < 3.5$
Milky Way stars	10 million	---

Mayall 4m
Kitt Peak, AZ

(Also talks by Michael Shubnell, Arjun Dey)

DESI at the NOAO Data Lab

datalab.noao.edu

Host DESI imaging Legacy Surveys (DECaLS, BASS/MzLS)

→ Databases (ls_dr3-dr7)

→ Images in NOAO Science Archive (raw + processed)

now!

Host a copy of DESI targets

→ Database for final, public set of targets

Host a copy of DESI redshifts

→ Database for public releases of redshift catalogs

→ Tools for spectra visualization/analysis

future

Create example Notebooks & workflows

Users can work with all data products

DESI at the NOAO Data Lab

datalab.noao.edu

Host DESI imaging Legacy Surveys (DECaLS, BASS/MzLS)

→ Databases (ls_dr3-dr7)

→ Images in NOAO Science Archive (raw + processed)

Host a copy of DESI targets

→ Database for final, public set of targets

Host a copy of DESI redshifts

→ Database for public releases of redshift catalogs

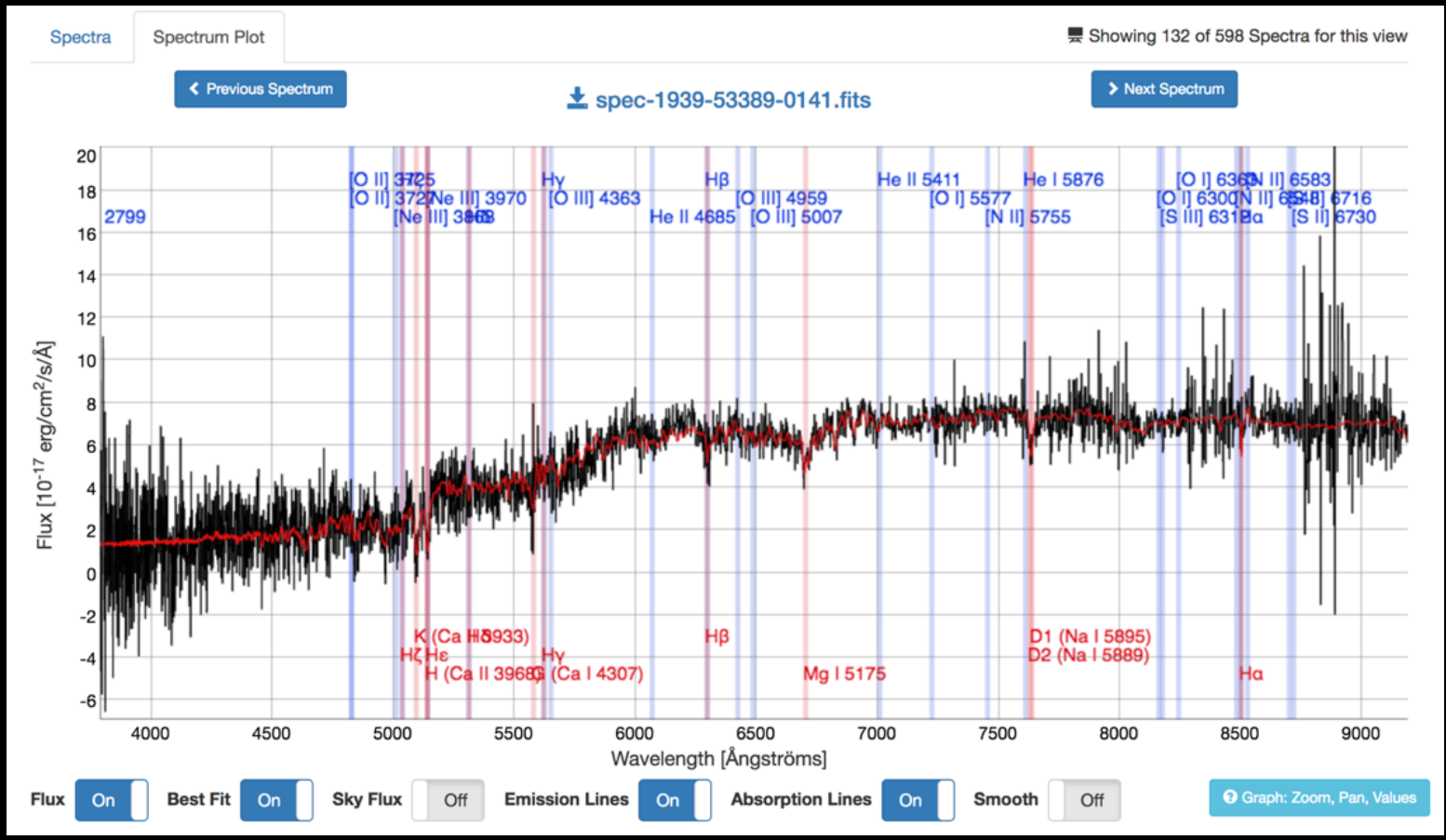
→ Tools for spectra visualization/analysis

Which one(s)?

Create example Notebooks & workflows

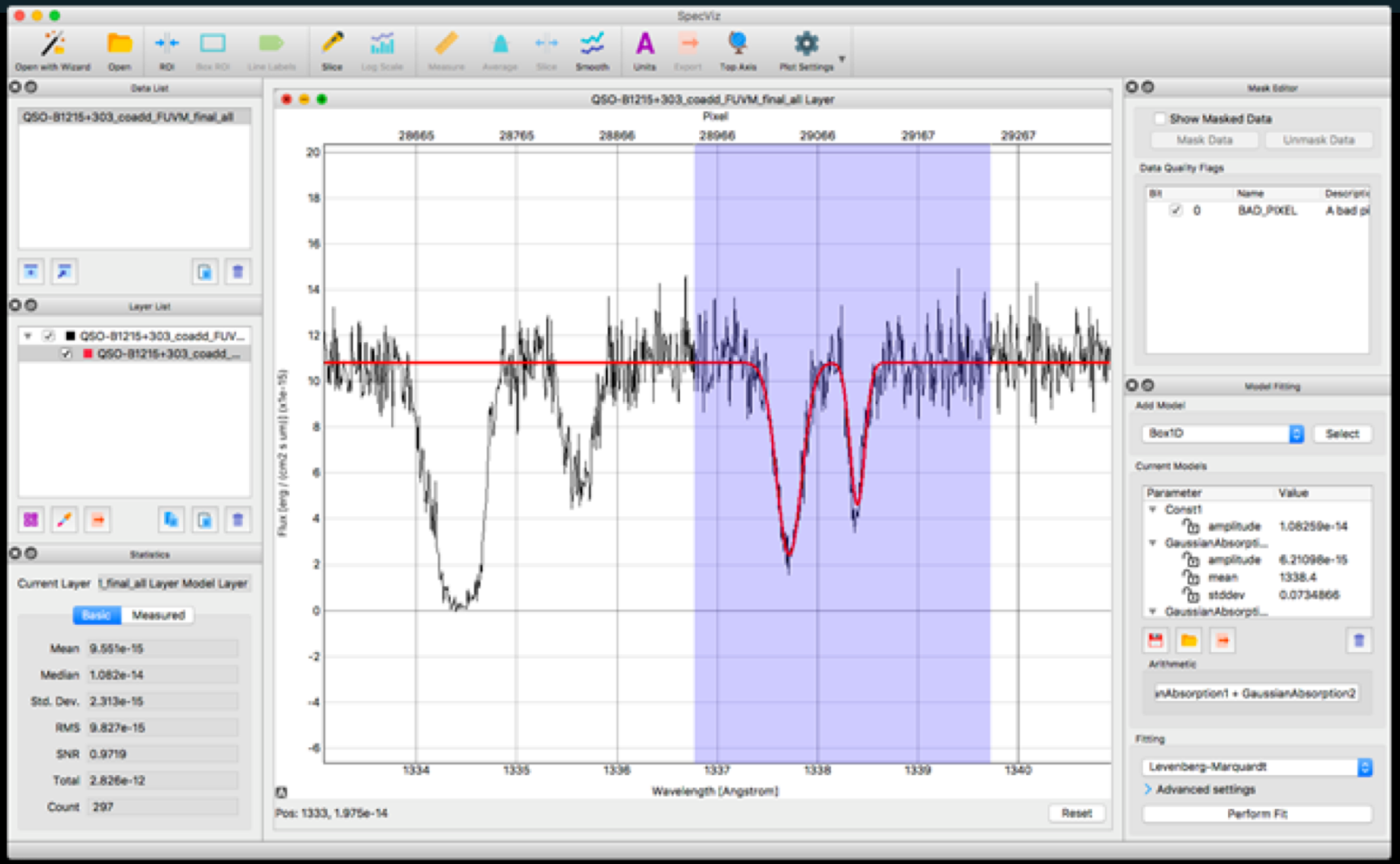
Users can work with all data products

Example: SDSS Science Archive Server



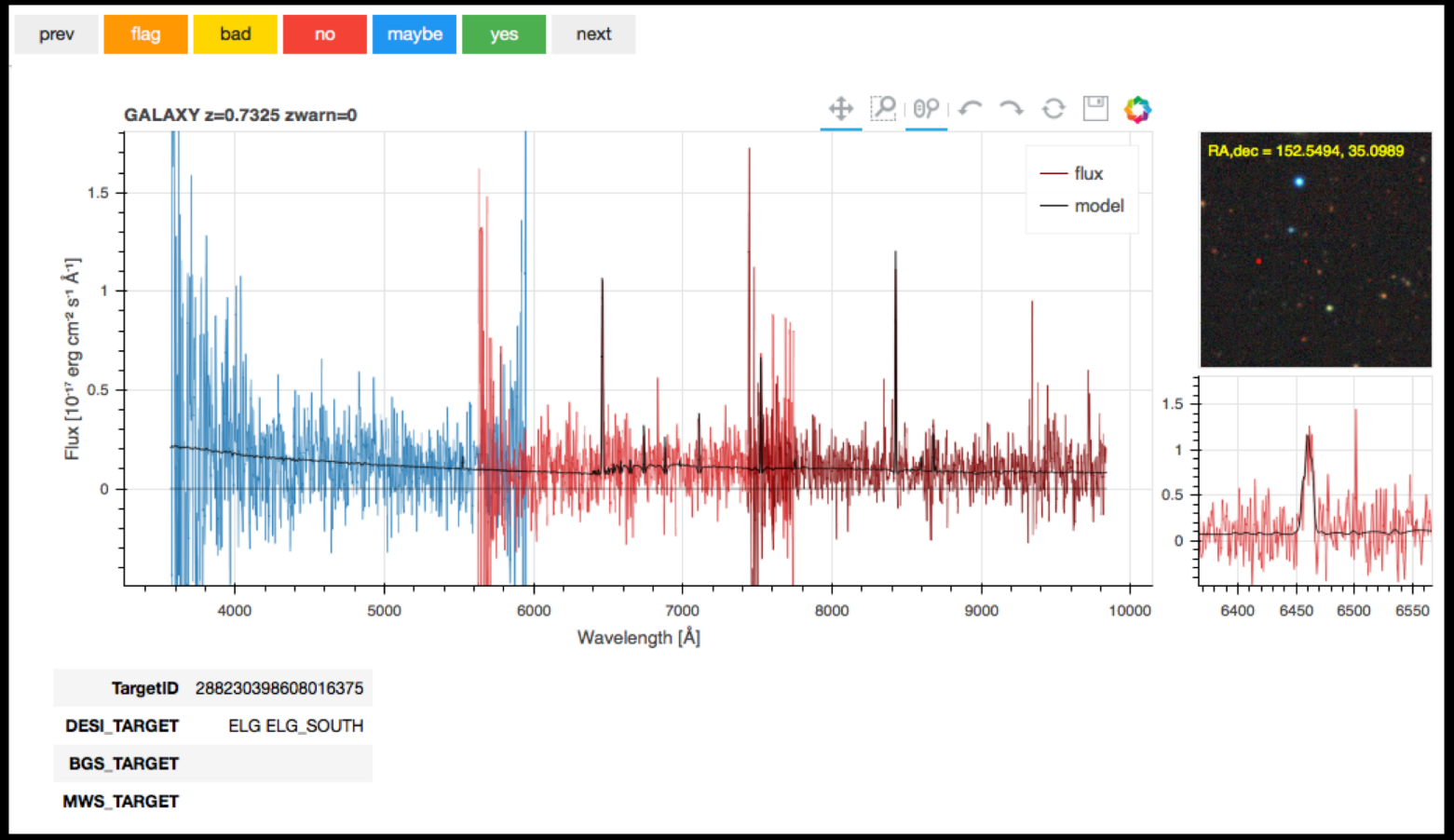
Interactive zoom/drag, optional display of spectral lines, best-fit template, sky, etc. → great for quick look in browser (but slow for large number)

Example: SpecViz (STSci)



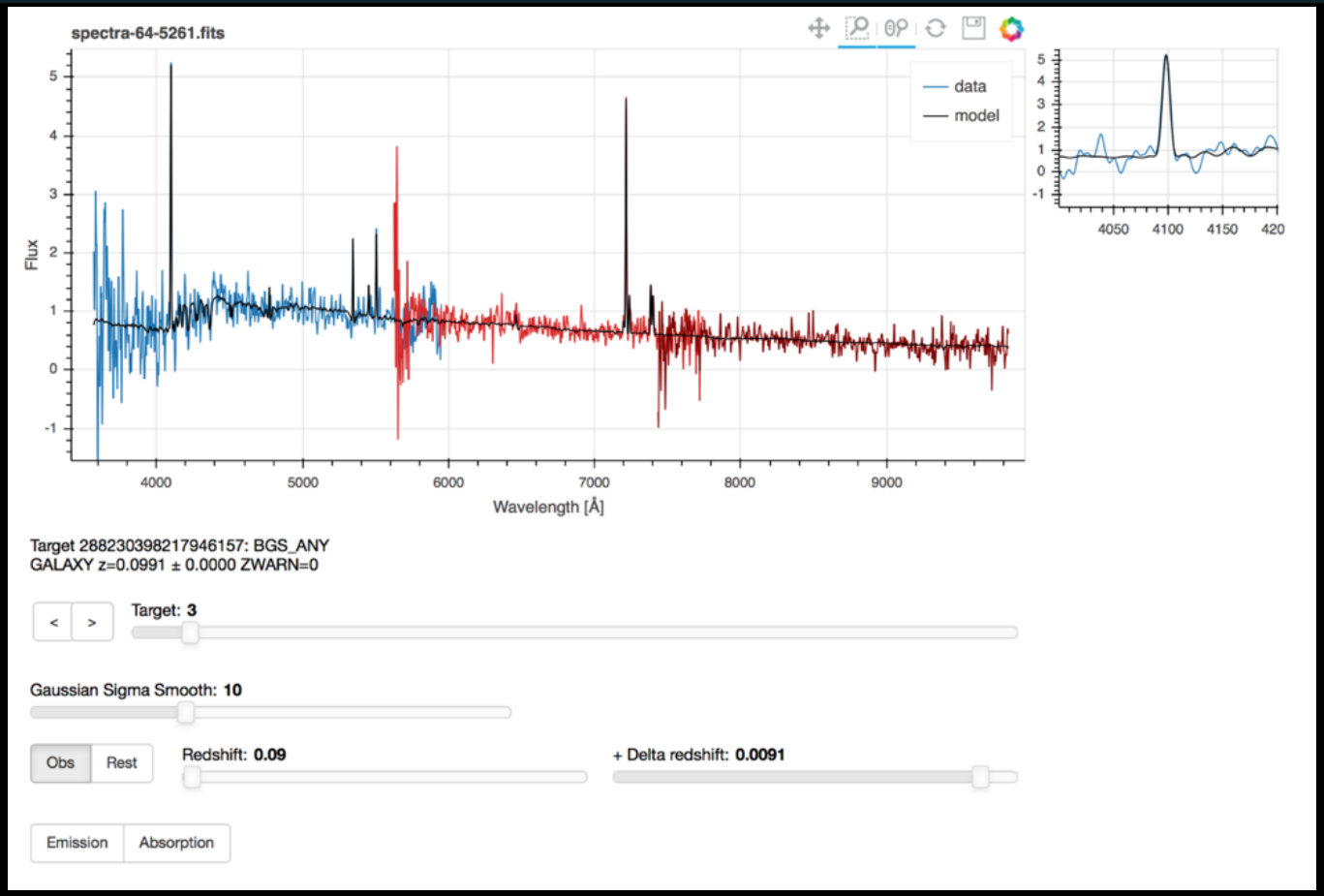
Interactive zoom/drag, fit of spectral features, etc. → stand-alone tool (slow for large number) + possible adaptation to Jupyter Notebook/Lab

Example: DESI *Inspector* prototype



Interactive functionality (Bokeh), "zoom box", contextual viewer w/ image
→ works in a Jupyter Notebook (but not in JupyterLab)

Example: DESI *prospect* prototype



Modified version of *inspector* → works in a JupyterLab + fast performance

In the end…

We wrote a “Spectral Viewer Requirements” document detailing what we need.

We will likely adapt an existing tool that comes close to meeting all the requirements, or maybe two different tools: (1) DESI-specific, and (2) general purpose

Recommendation for MSE: spend time defining the data model, software, analysis tools needs early on – if possible in coordination with other wide-field spectroscopy effort

Goals
(in the context of the future data ecosystem)

Easy access to data for entire astronomy community

→ Databases: Tables, Images, Spectra

(data flood from many projects: the ones that get used the most for discoveries will make a long-lasting impression)

User-friendly yet powerful analysis tools

→ Quick start analysis

→ Automated & sophisticated workflows

Interactive interface with advanced visualization

→ connected exploration & analysis

Advanced & strategic thinking required!

Wide field spectroscopy
in the future data ecosystem

We're not ready…

… but we can be!

Maunakea Spectroscopic Explorer

National Optical Astronomy Observatory
Cerro Tololo Inter-American Observatory
Kitt Peak National Observatory
Community Science and Data Center

AURA