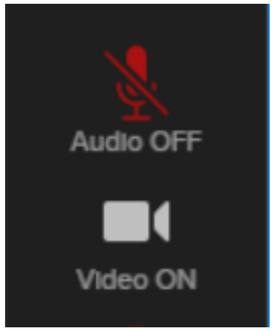
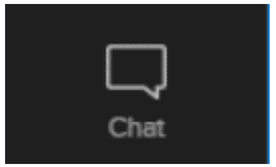
Welcome to the SGCI Webinar!

- We will be starting shortly.
- Your audio has been muted, and you are encouraged to turn off your video during the presentation.
  - Controls for these are near the bottom of the right-side control panel for BlueJeans.



- You may submit questions at any time using Chat, and the moderator will share them with the presenter when appropriate.



- This presentation will be recorded and slides will be posted.

SGCI's week-long **Science Gateways Bootcamp** in October teaches strategies for successful gateway development & sustainability.

Apply by Friday 7/28: https://sciencegateways.org/bootcamp

A few brief words about the Science Gateways Community Institute (SGCI)

Our goal: To facilitate community *sharing of experiences, technologies, and practices* at little or no cost to community members through *NSF-funded, online and in-person resources and services*

Incubator: Learn best practices from our consultants or Bootcamp.

Extended Developer Support: Get direct, custom development help.

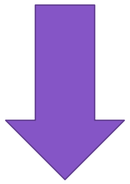
Scientific Software Collaborative: Find gateways or software components (or promote your own).

Community Engagement & Exchange: Engage with and learn from the gateways community.

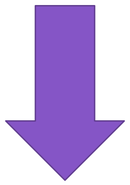
Workforce Development: Build your professional career as a student or young professional.

A quick favor at the end of this webinar…

NSF gives money to SGCI.



SGCI gives you a free webinar.



Could you give SGCI
30 seconds of feedback?

jupyter

A Gateway for Scientific Collaboration and Education

The Project Jupyter Team

Carol Willing, Cal Poly
Brian Granger, Cal Poly
Fernando Perez, LBNL/Berkeley
Min Ragan-Kelley, Simula

The Larger Jupyter Team

@ProjectJupyter on Twitter

July 20, 2017

Proud member of the Jupyter community

Carol Willing

- Steering Council, Project Jupyter
- Core Developer, Project Jupyter
- Software Engineer, Cal Poly SLO
- Director, Python Software Foundation
- Core Developer, CPython
- Geek in Residence, Fab Lab San Diego

@willingcarol

Agenda

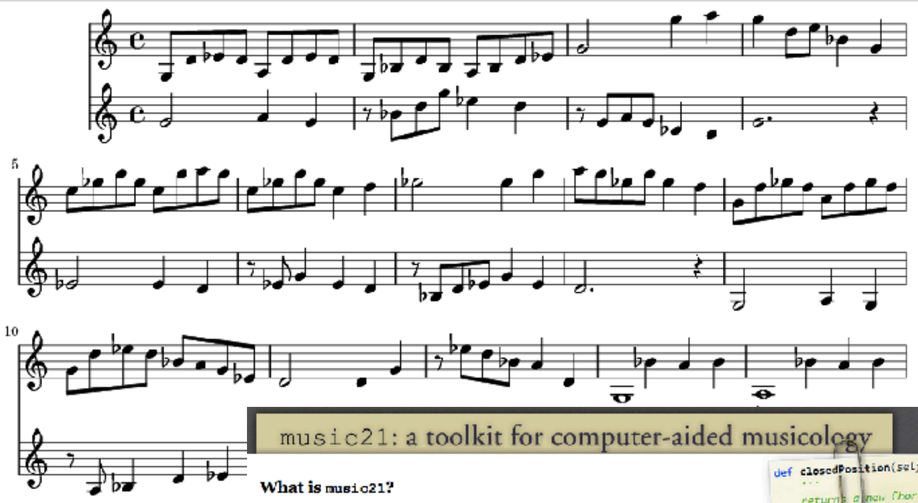
- Gateways
- Jupyter Notebook
- JupyterHub
- JupyterLab
- Next steps

2013 - Engage users

Convert from kern format to MusicXML

```
In [12]: c = converter.parse('/Users/carol/Downloads/duet/edokomuri.krn')
         c.show()
```

Out[12]:



music21: a toolkit for computer-aided musicology

What is music21?

music21 is a set of tools for helping scholars and other active listeners answer questions about music quickly and simply. If you've ever asked yourself a question like, "I wonder how often Bach does *this*" or "I wish I knew which band was the first to use these chords in this order," or "I'll bet we'd know more about Renaissance counterpoint (or Indian ragas or post-tonal pitch structures or the form of minuets) if I could write a program to automatically write more of them," then music21 can help you with your work.

How simple is music21 to use?

Extremely. After starting Python and typing "from music21 import *" you can do all of these things with only a single line of music21 code:

Display a short melody in musical notation:

```
converter.parse("tinynotation: 3/4 c4 d8 f g16 a g f#").show()
```

Print the twelve-tone matrix for a tone row (in this case the opening of Schoenberg's Fourth String Quartet):

```
print (serial.rowToMatrix([2,1,0,10,8,5,4,7,6,9,3,11]))
```

or since all the end-Viennese school rows are already available as objects, you can type:

```
print (serial.getHotMatrix('BrnoSchoenbergOp37')).matrix() )
```

Convert a file from Humdrum's **kern data format to MusicXML, for editing in Finale or Sibelius:

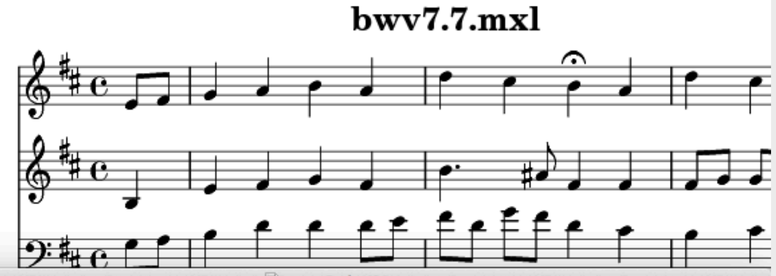
```
converter.parse('/users/cuthbert/docs/composition.krn').write('musicxml')
```

```
def closedPosition(self):
    '''
    returns a new Chord object with
    >>> chord1 = Chord(["C4", "G5",
    >>> chord2 = chord1.closedPositio
    >>> print(chord2.tiny.value)
    <clef: e' g'>4
    '''
    newChord = copy.deepcopy(self)
    tempChordNotes = newChord.pitches
    chordBassPS = self.bass().ps
    for thisPitch in tempChordNotes:
        while thisPitch.ps > chordBas
            thisPitch.octave = thisP
    newChord.pitches = tempChordNote
```

- Get Started with music21
- Browse the music21 documentation
- Download music21 from Google Code
- Get our latest news and updates at the music21 blog
- Read the Frequently Asked Questions list
- Sign up for the music21list mailing list through Google Groups.

```
In [10]: sBach = corpus.parse('bach/bwv7.7')
         sBach.show()
```

Out[10]:



tmpOph_K4.txt

```
Movement Name: bwv7.7.mxl
```

Michael Scott Cuthbert (cuthbert [at] mit.edu) is Associate Professor of Music and Homer A. Burnell Career Development Professor at M.I.T.

2014 - Break down barriers to entry

Intro to Python

San Diego Python

- Start with a proven curriculum

  http://pyvideo.org/pycon-us-2013/a-hands-on-introduction-to-python-for-beginning-p.html

- Hands on to engage students
- Takeaway notebooks reduce student stress

  https://github.com/pythonsd/intro-to-python

2014 - Build a community

Andrea Zonca: notes about Python, high performance computing, data analysis

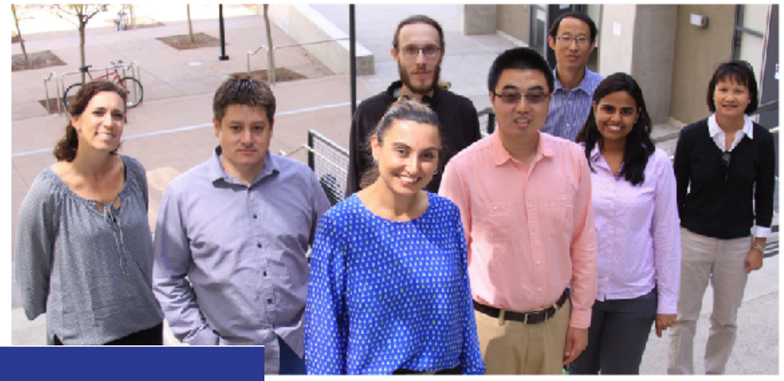
Posts:

How to create pull requests on Github	[FRI 30 JUNE 2017]
Deploy Jupyterhub on a Supercomputer with SSH Authentication	[TUE 16 MAY 2017]
Configure Globus on your local machine for GridFTP with XSEDE authentication	[WED 19 APRIL 2017]
Sample deployment of Jupyterhub in HPC on SDSC Comet	[SUN 26 FEBRUARY 2017]
Customize your Python environment in Jupyterhub	[FRI 24 FEBRUARY 2017]
Automated deployment of Jupyterhub with Ansible	[FRI 03 FEBRUARY 2017]
How to publish your research software to Github	[WED 01 FEBRUARY 2017]
Run Ubuntu in HPC with Singularity	[FRI 13 JANUARY 2017]
Jupyterhub Docker Spawner with GPU support	[WED 12 OCTOBER 2016]
Jupyterhub deployment on multiple nodes with Docker Swarm	[TUE 24 MAY 2016]
Quick Jupyterhub deployment for workshops with pre-built image	[THU 28 APRIL 2016]
Deploy Jupyterhub on a Virtual Machine for a Workshop	[SAT 16 APRIL 2016]
Use your own Python installation (kernel) in Jupyterhub	[MON 05 OCTOBER 2015]
IPython/Jupyter notebook setup on NERSC Edison	[THU 24 SEPTEMBER 2015]
IPython/Jupyter notebook setup on SDSC Comet	[THU 17 SEPTEMBER 2015]
Run Jupyterhub on a Supercomputer	[THU 02 APRIL 2015]
Accelerate groupby operation on pixels with Numba	[TUE 24 MARCH 2015]
Software Carpentry setup for Chromebook	[TUE 10 FEBRUARY 2015]
Zero based indexing	[WED 22 OCTOBER 2014]
Write unit tests as cells of IPython notebooks	[TUE 30 SEPTEMBER 2014]

WOrDS Workflows for Data Science Center of Excellence

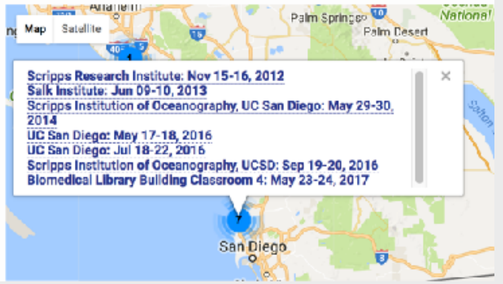
Home | About | Team | Projects | Research | Publications | Events | Contact | Login

Team



software carpentry

Past Workshops



Credits: http://zonca.org/ https://words.sdsc.edu/team https://software-carpentry.org

Serve users and foster collaboration

2017 and beyond

Incubator

Learn best practices from our consultants or Bootcamp

Extended Developer Support

Get direct, custom development help

Scientific Software Collaborative

Find gateways or software components (or promote your own)

Community Engagement & Exchange

Engage with and learn from the gateways community

Workforce Development

Build your professional career as a student or young professional

Credit: http://sciencegateways.org/


IPython
Parallel
Terminal

> **Project Jupyter** serves not only the academic and scientific communities but also a much broader constituency of data scientists in research, education, industry and journalism…

- Fernando Pérez
UC Berkeley

“…we see uses of our tools that range from **high school education** in programming to the nation’s **supercomputing** facilities and the **leaders of the tech industry**.

- Fernando Pérez
UC Berkeley

“

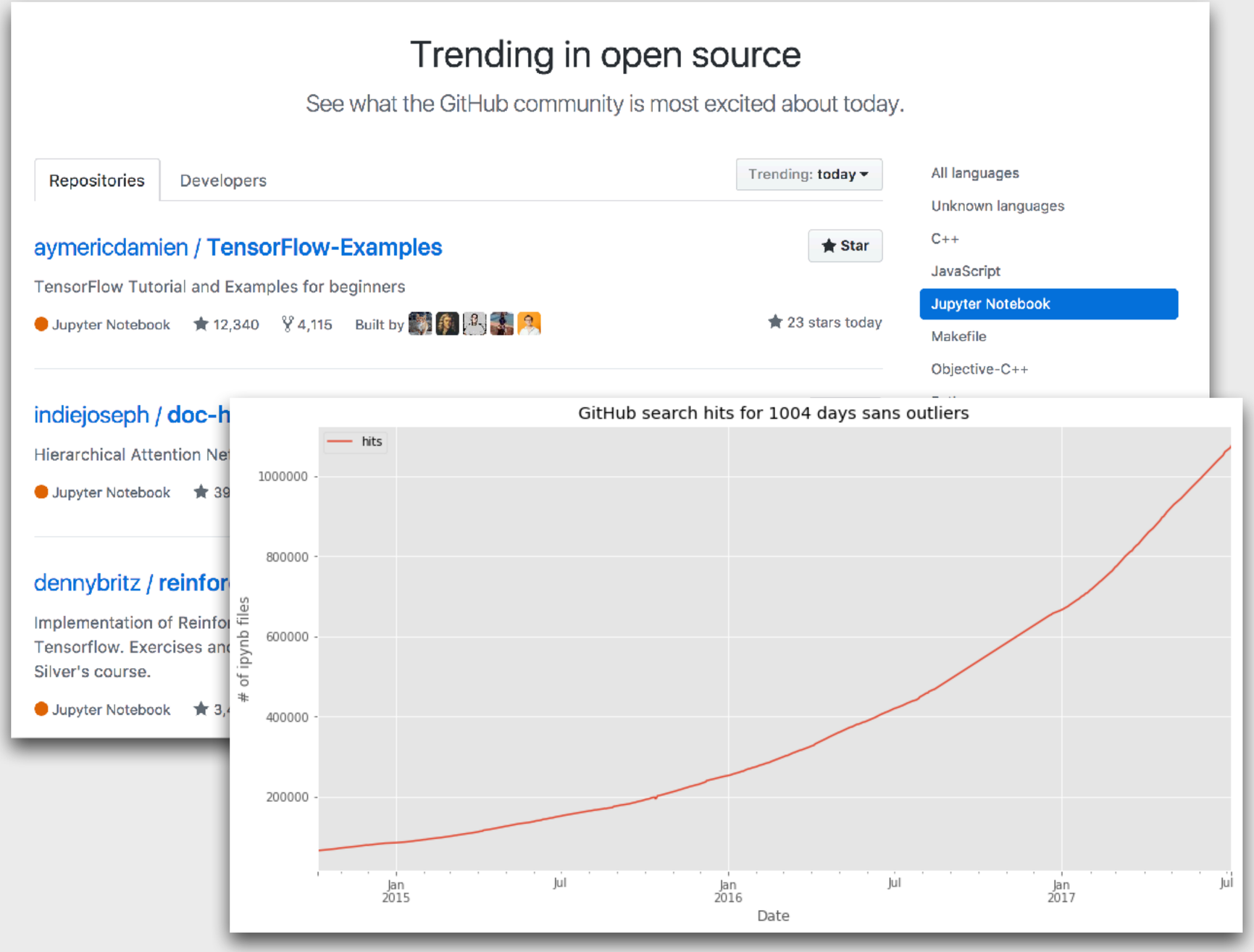
More than **a million people** are currently using Jupyter for everything from…

-Prof. Brian Granger
Cal Poly

>6M Users

https://github.com/jupyter/design/blob/master/surveys/2015-notebook-ux/analysis/report_dashboard.ipynb

Over 1M Notebooks on GitHub



https://github.com/trending/jupyter-notebook

> “…analyzing massive gene sequencing **datasets** to processing **images** from the Hubble Space Telescope and developing **models** of financial markets.

-Prof. Brian Granger
Cal Poly

Scaling globally

Credit: http://pythonineducation.org/

> We are excited by the potential of Project Jupyter to **reach even wider audiences** and to contribute to **increased cross-disciplinary collaboration** in the sciences.

-Betsy Fader
Helmsley Charitable Trust

“

Jupyter Notebook… will enable data exploration, visualization, and analysis in a way that encourages sound science and speeds progress.

-Chris Mentzel

The Gordon and Betty Moore Foundation

Agenda

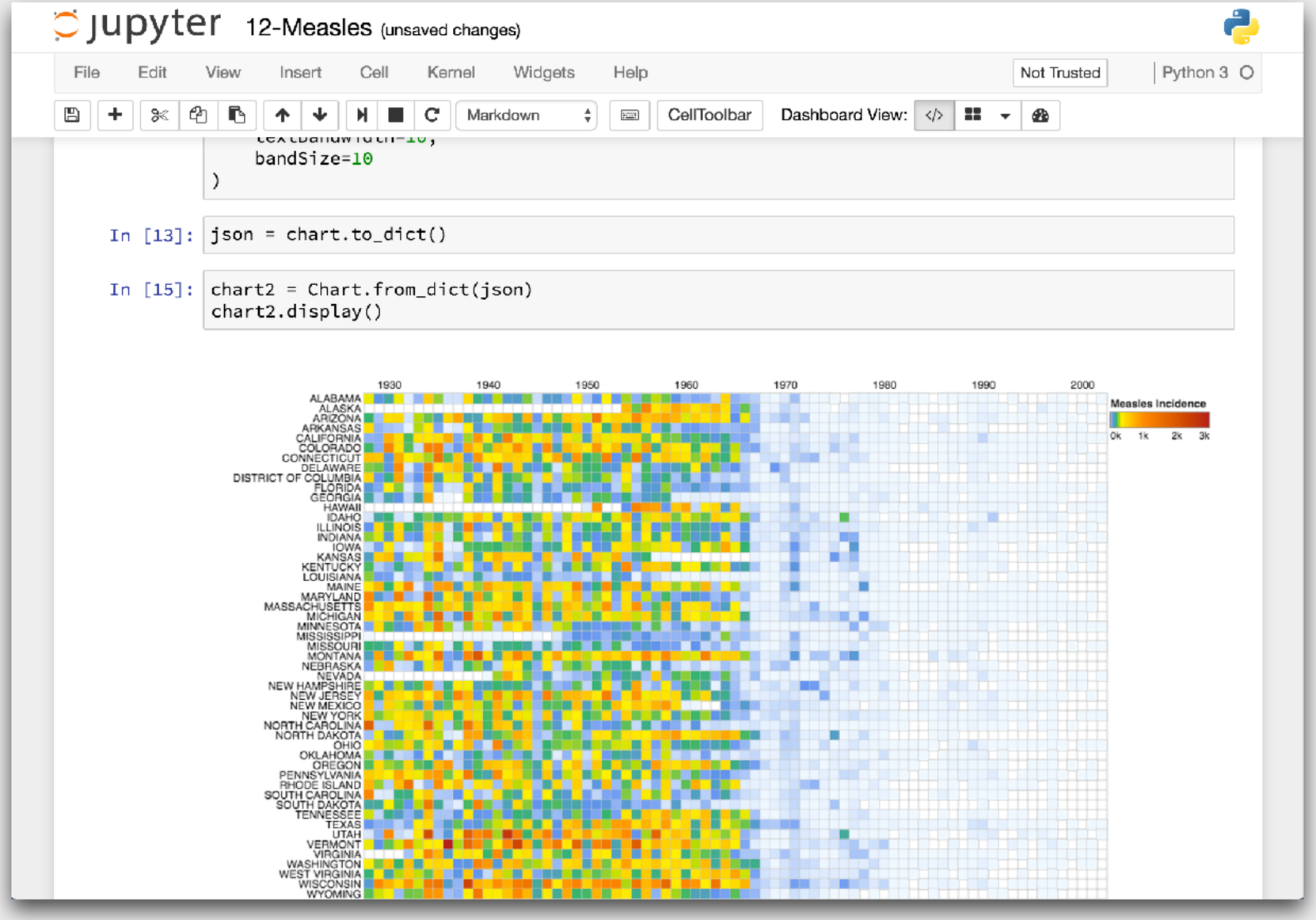
- Gateways
- Jupyter Notebook
- JupyterHub
- JupyterLab
- Next steps

Jupyter Notebook

Interactive, Exploratory, Reproducible

- Interactive, browser-based computing environment
- Exploratory data science, ML, visualization, analysis, stats
- Reproducible document format:
  - Code
  - Narrative text (markdown)
  - Equations (LaTeX)
  - Images, visualizations
- Over 50 programming languages
- Everything open-source (BSD license)

Jupyter Notebook



A Jupyter Notebook document with a visualization of measles data.

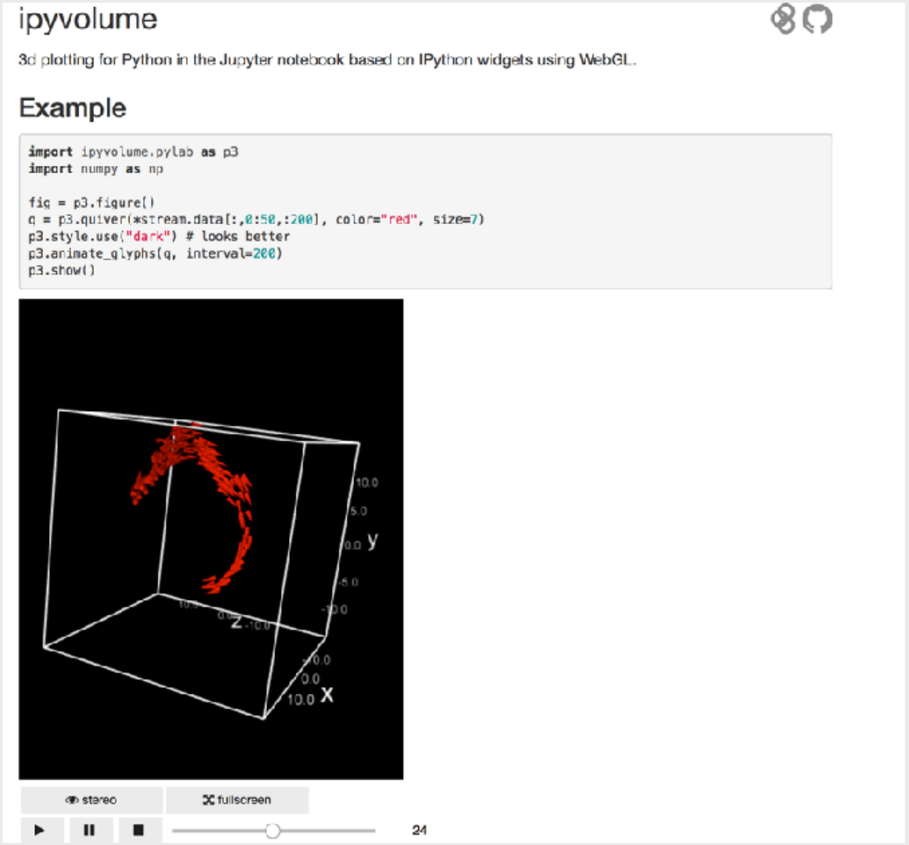
Interactive Documentation

Engaging User Content

Rapid “what if” scenarios

ipywidgets

- **Docs** https://ipywidgets.readthedocs.io
- **Website** http://jupyter.org/widgets.html
- **Blog 6.0 release** https://blog.jupyter.org/2017/03/01/ipywidgets-6-release/
- **cookiecutter** to simplify creating new widgets

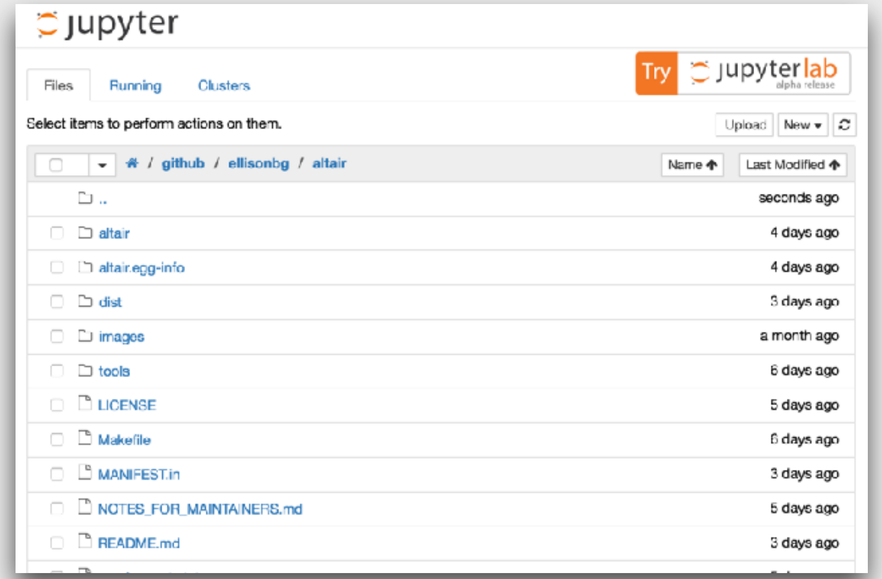


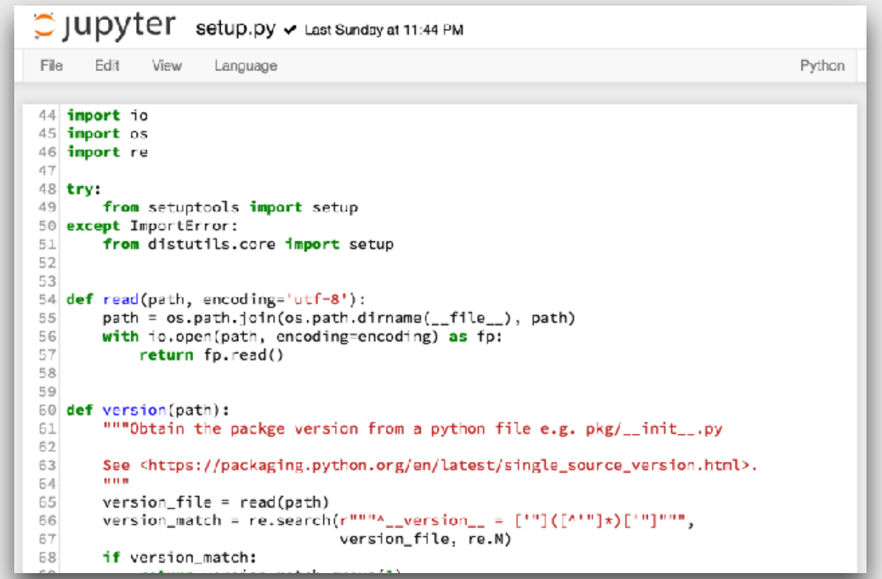
Pushing the boundaries

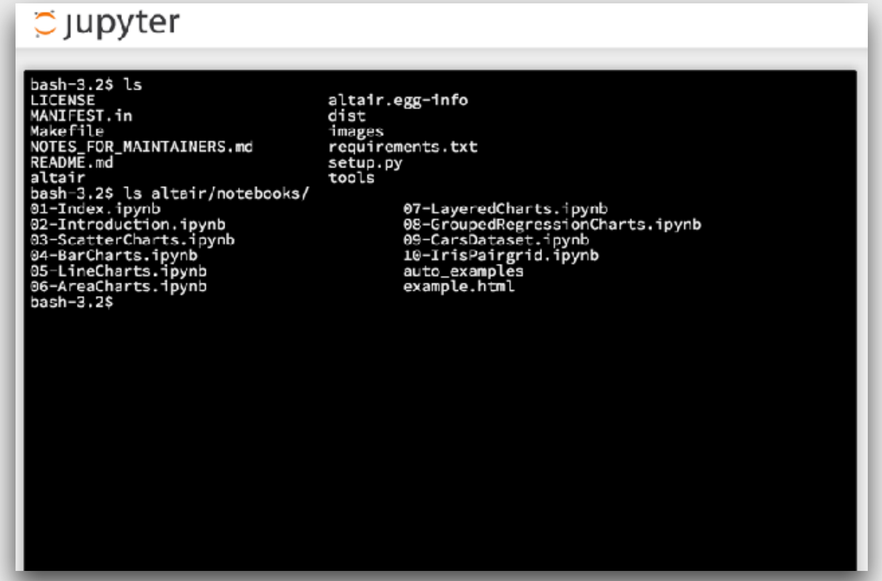
ipyvolume

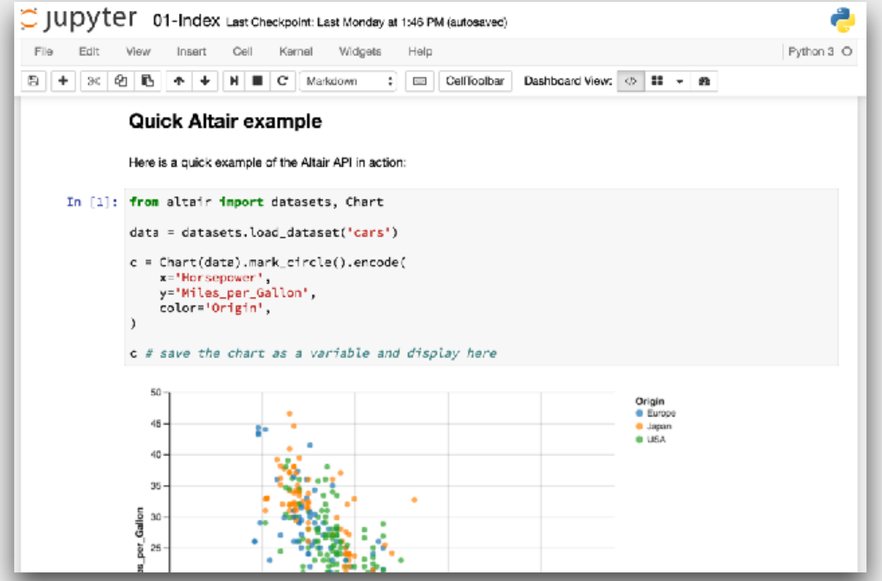
- **3D interactivity** in notebooks
- **Innovation** by Maarten Breddels and team
- **Documentation** engages and demonstrates
- **Try and enjoy** at ipyvolume.readthedocs.io

Classic Jupyter: More Than Just Notebooks





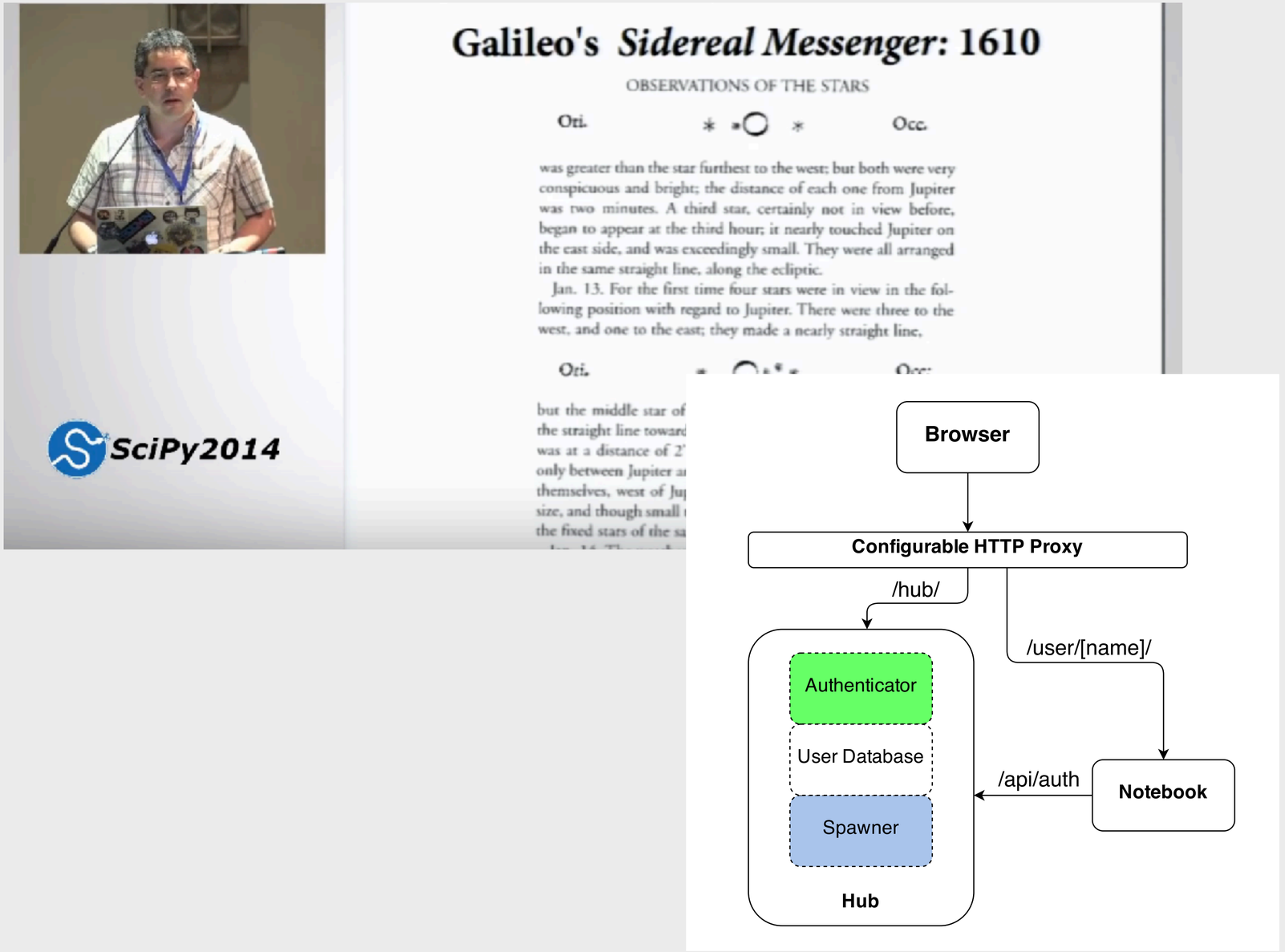




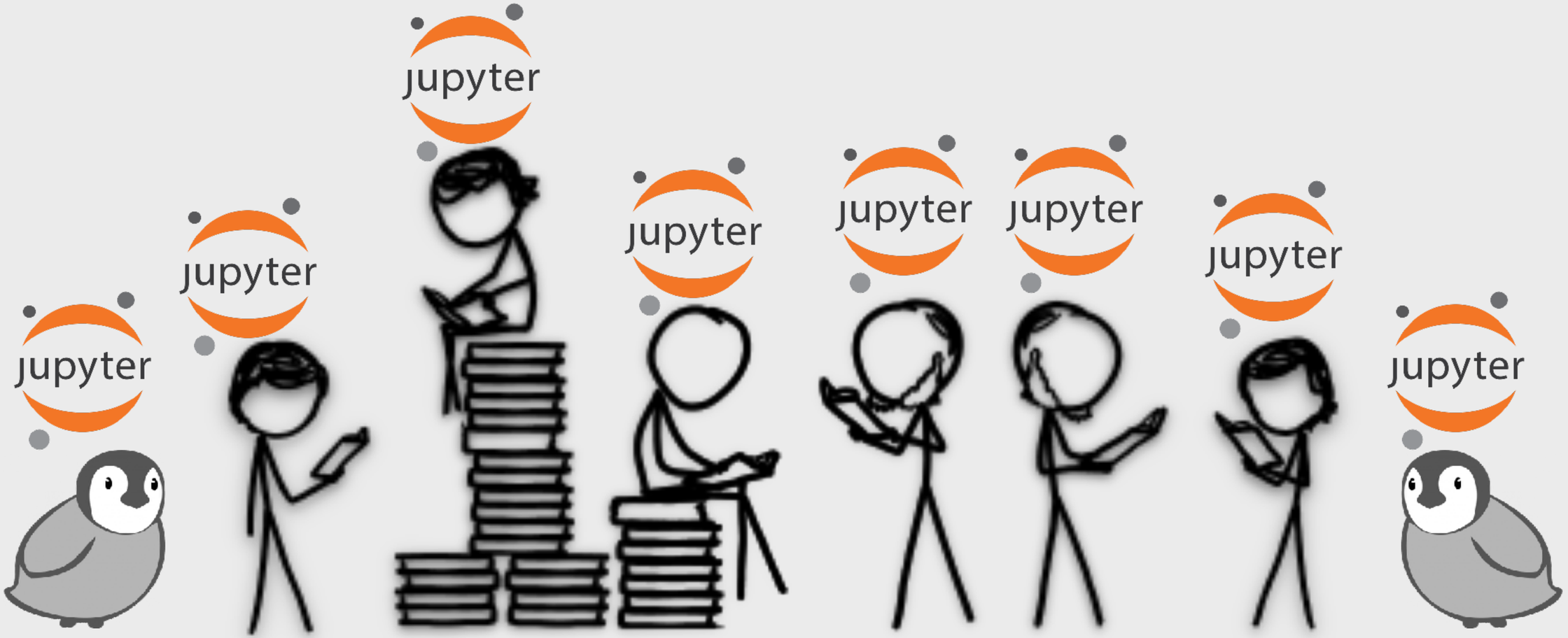
Agenda

- Gateways
- Jupyter Notebook
- JupyterHub
- JupyterLab
- Next steps

Jupyter for Science and Data Science



JupyterHub


jupyter
jupyter
jupyter
jupyter
jupyter
jupyter
jupyter
jupyter

Exploration and prototyping

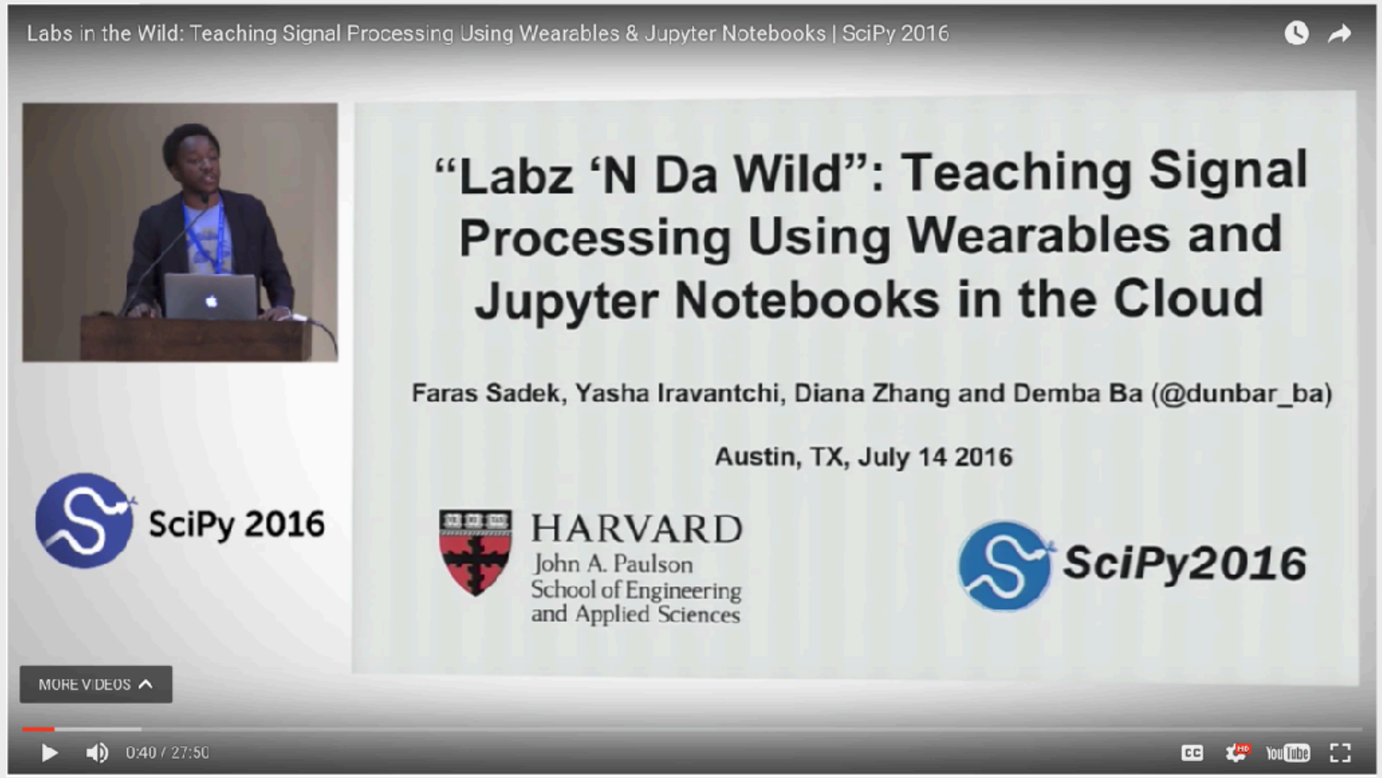
Teaching Signal Processing using Wearables and Jupyter Notebooks

Dr. Demba Ba

- Exploration and experimentation

  http://pyvideo.org/scipy-2016/labs-in-the-wild-teaching-signal-processing-using-wearables-jupyter-notebooks-scipy-2016.html

- Physical media with wearables and electronics
- Real world, self-directed projects



Visualize and communicate

Python for Geosciences

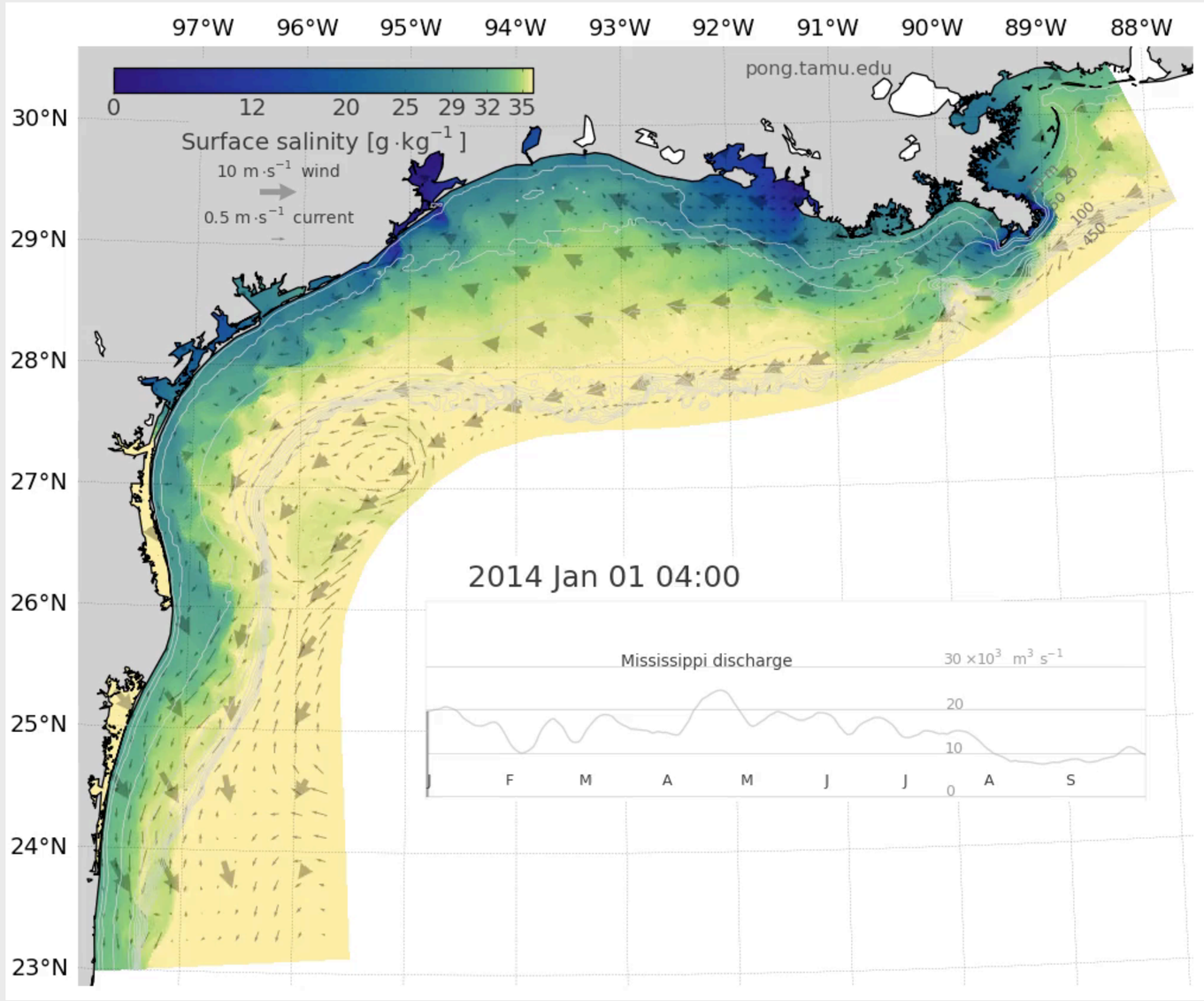
Dr. Kristen Thyng

- Feedback and communication with students using nbgrader

  http://kristenthyng.com/blog/2016/09/07/jupyterhub+nbgrader/

- Progression to complex examples and tasks

  https://github.com/kthyng/python4geosciences



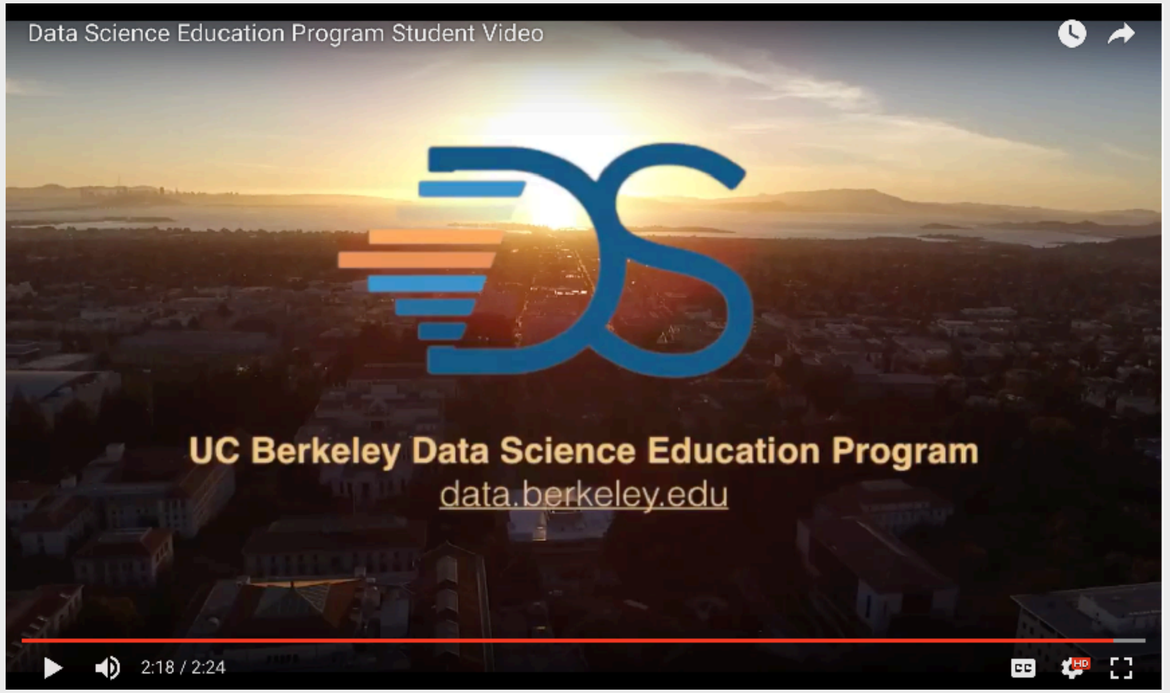
Scale learning with research tools

Berkeley Data Science Data8

UC Berkeley

- Campus wide curriculum
- Cross-discipline
- Zero to JupyterHub with Kubernetes

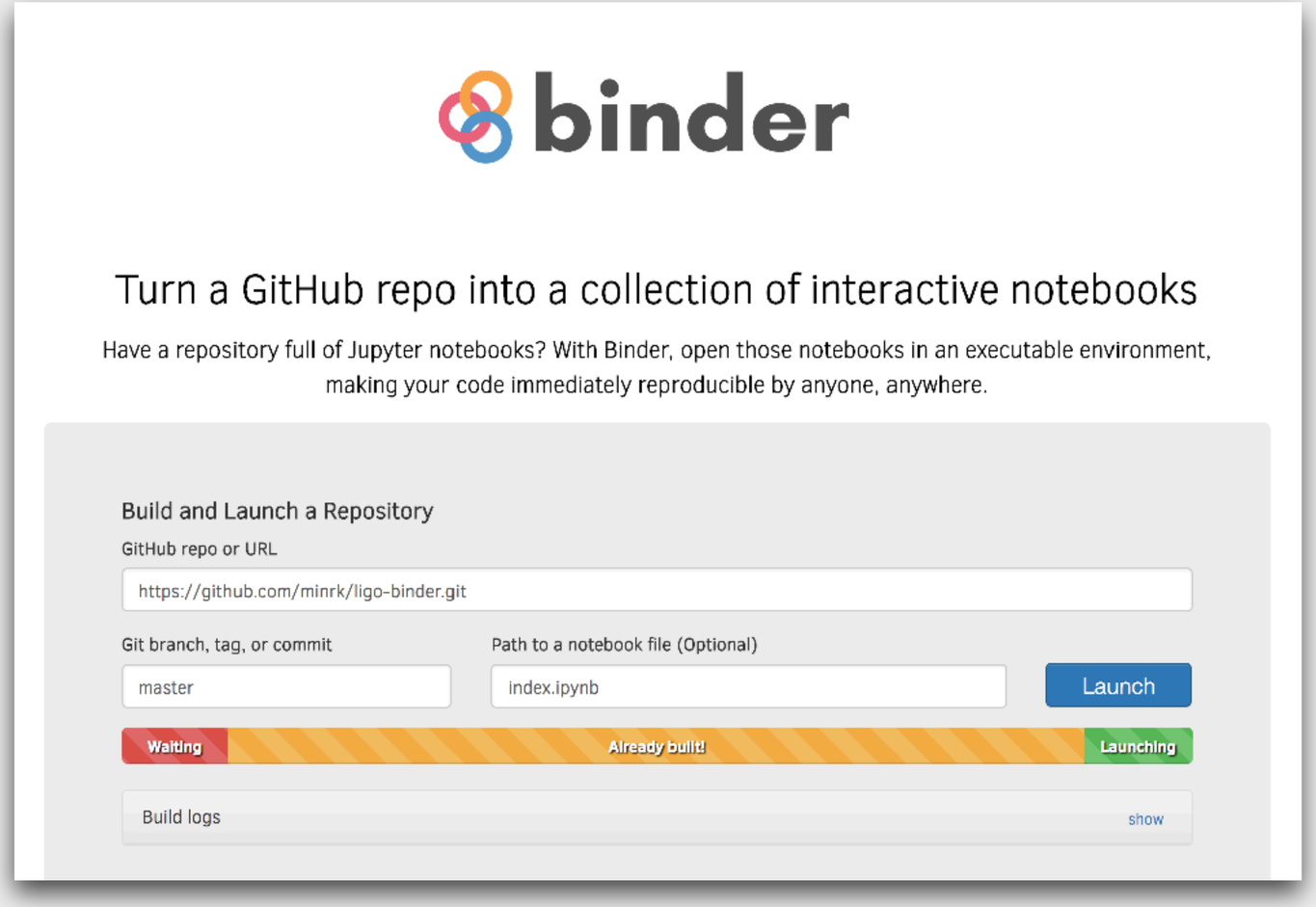
https://zero-to-jupyterhub.readthedocs.io



http://denero.org/data-8-in-spring-2017.html
https://github.com/data-8/jupyterhub-k8s
http://data8.org/
http://data.berkeley.edu/
http://data.berkeley.edu/about/videos

Live Code on Binder

LIGO Binder



https://losc.ligo.org/tutorials/

https://beta.mybinder.org/v2/gh/minrk/ligo-binder/master?filepath=index.ipynb

Enabling Reproducible Science

https://losc.ligo.org/about/

JupyterHub

Where we are

0.7 - 12/2016

- introduce **Services**
- Anything that can talk to the Hub's API that's not a **User**
- **Managed Service:** A process started by the Hub
- **External Service:** Anything not started by the Hub (may or may not be a process)

Where we are

0.7 Services can...

- run a web service at /services/:service-name
- authenticate requests with the **Hub** via **HubAuth**
- talk to the **Hub API** with their **API token**(s)

Where we are

0.7 Services are for...

- interacting with the **Hub**
- nbgrader formgrader
- culling idle servers
- sharing files
- shared notebook server(s)
- nbviewer

Where we are going

JupyterHub

0.8

Where we are going

0.8

- abstract **Proxy** API
  - define spec and Python API for Hub's proxy needs
  - Better support nginx, kubernetes proxies
  - Requires moving activity tracking to single-user servers (done in notebook 5.0)

Where we are going

0.8

- multiple **servers** per user
  - useful when single **Hub** exposes a variety of computational resources (clusters)
  - servers can have different configurations (different Spawners?)
  - need to keep common single server-per-user case well supported, to avoid overcomplicating things
  - contributions started by Christian Barra

Where we are going

OAuth

- JupyterHub as **OAuth provider**
  - removes need for complicated cookie management by the Hub
  - Will be needed as number of endpoints for which users are authorized grows (shared servers for collaboration)

Where we are going

HubShare

- **Service** for sharing
  - unit of sharing: directory
  - push/pull model
  - simple REST spec (possibly WebDAV)
  - share with individuals, groups
  - target use case: nbgrader assignments

https://github.com/jupyterhub/hubshare

Agenda

- Gateways
- Jupyter Notebook
- JupyterHub
- JupyterLab
- Next steps

Building Blocks

File Browser

Notebooks

Terminal

Text Editor

Kernels

Output

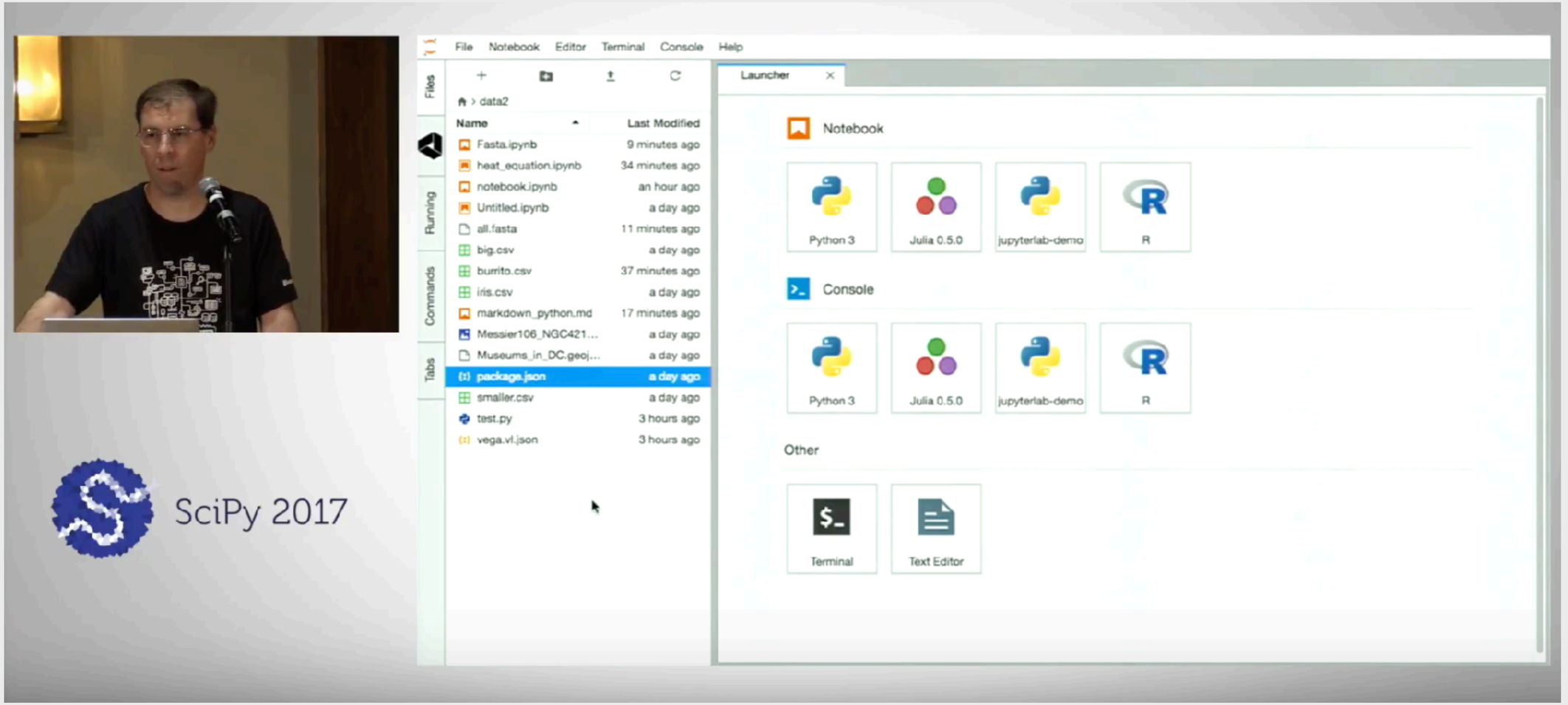
Introducing JupyterLab:
The Evolution of the Jupyter Notebook (almost beta)

The JupyterLab Team

Chris Colbert, Continuum
Steven Silvester, Continuum
Afshin Darian, Continuum
Jason Grout, Bloomberg
Brian Granger, Cal Poly
Grant Nestor, Cal Poly
Cameron Oelsen, Cal Poly
Fernando Perez, LBNL/Berkeley
Ian Rose, Berkeley
Cal Poly Interns
The Larger Jupyter Team

@jupyterlab on GitHub
@ProjectJupyter on Twitter

A Whirlwind Tour of JupyterLab



Learn more about

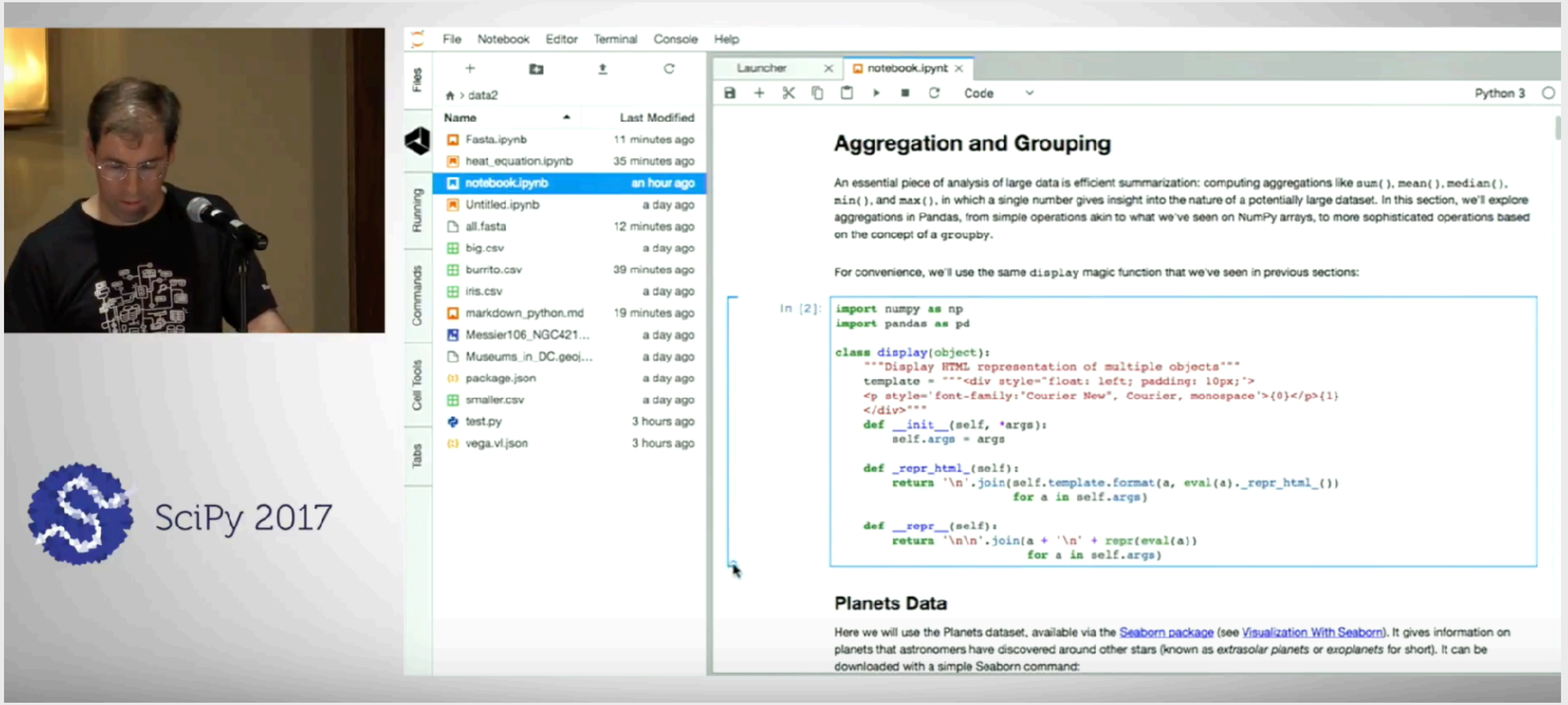
JupyterLab

PyData Seattle: https://www.youtube.com/watch?v=tHZT_mpNRcY

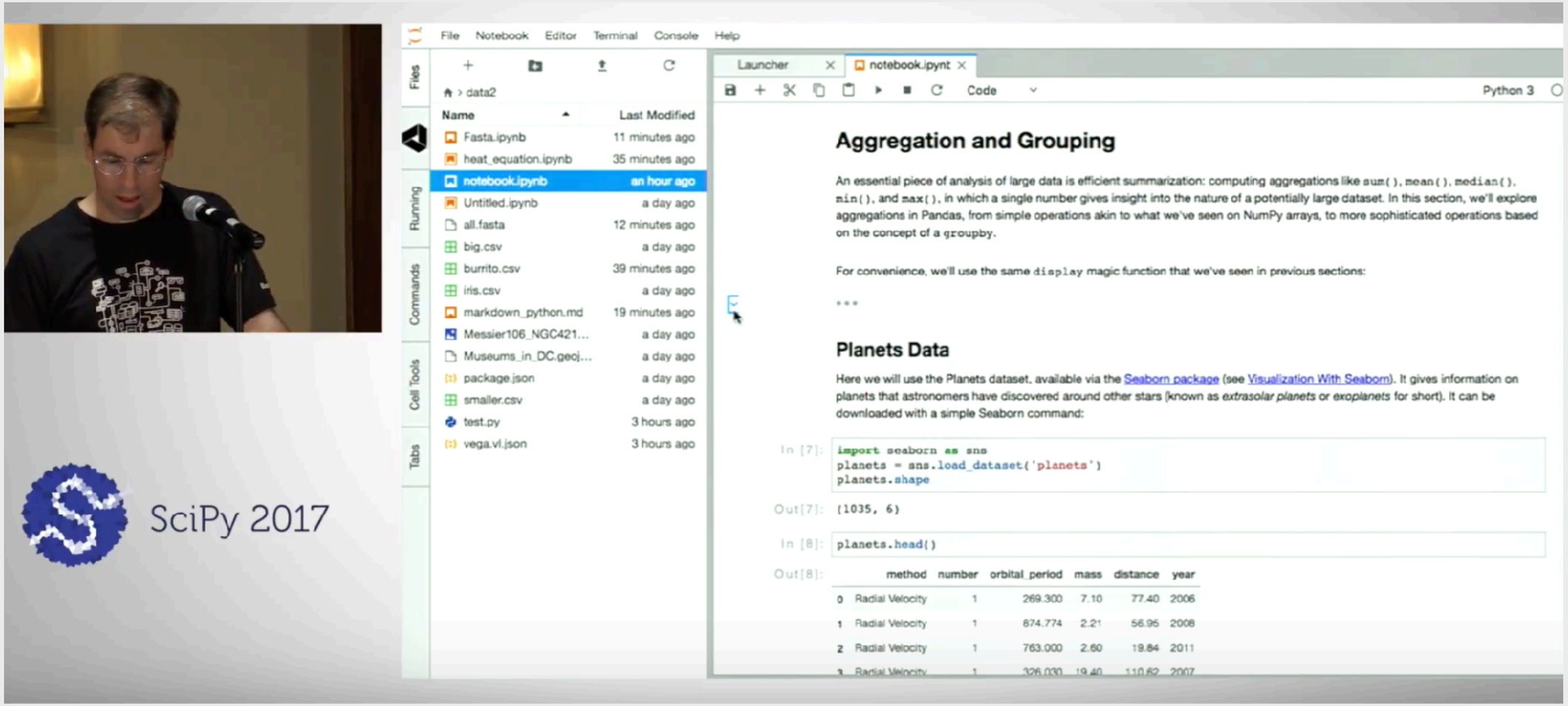
Demo from SciPy 2017: https://youtu.be/X8zPuBu22Y4?t=44m50s

http://pyvideo.org/pydata-dc-2016/jupyterlab-building-blocks-for-interactive-computing.html

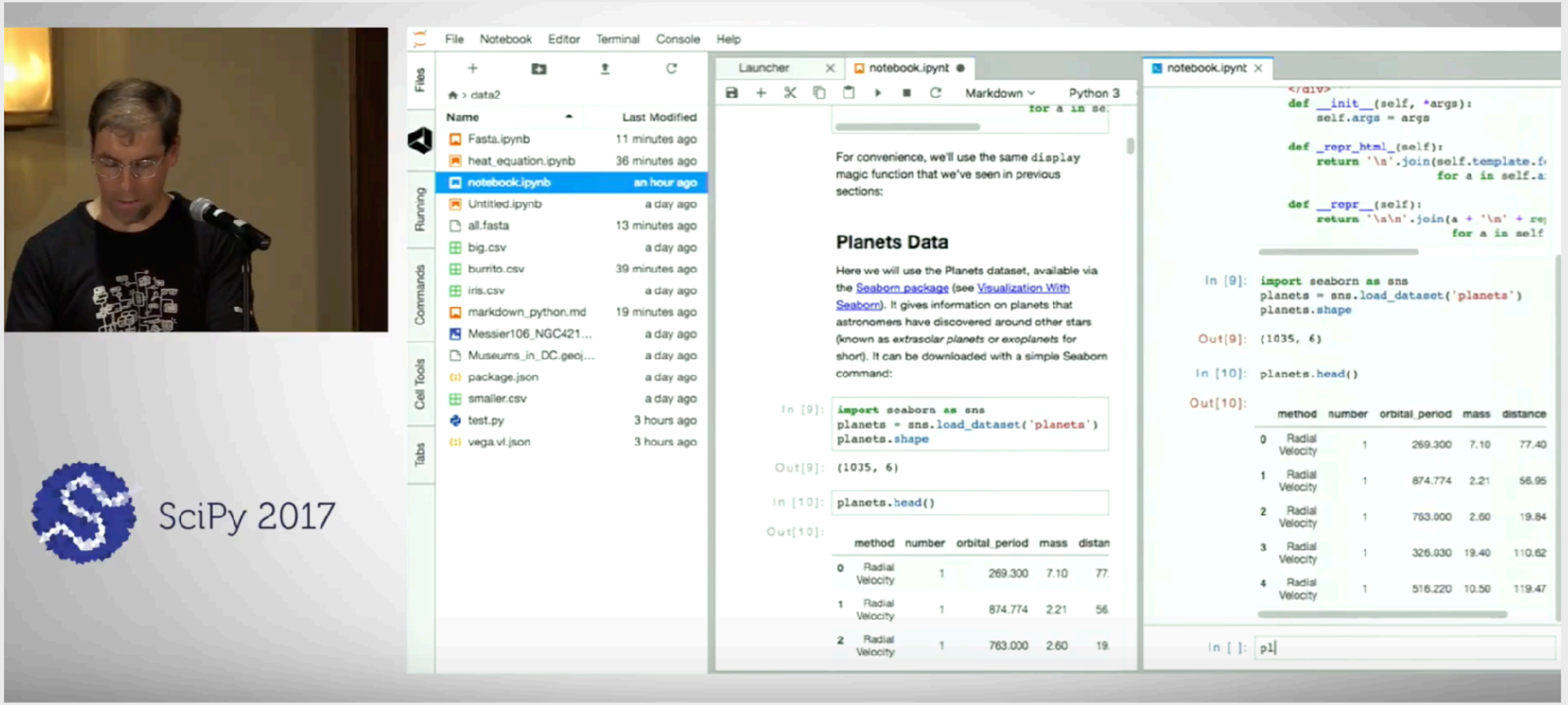
New implementation of the notebook



Collapsible cells and draggable cells



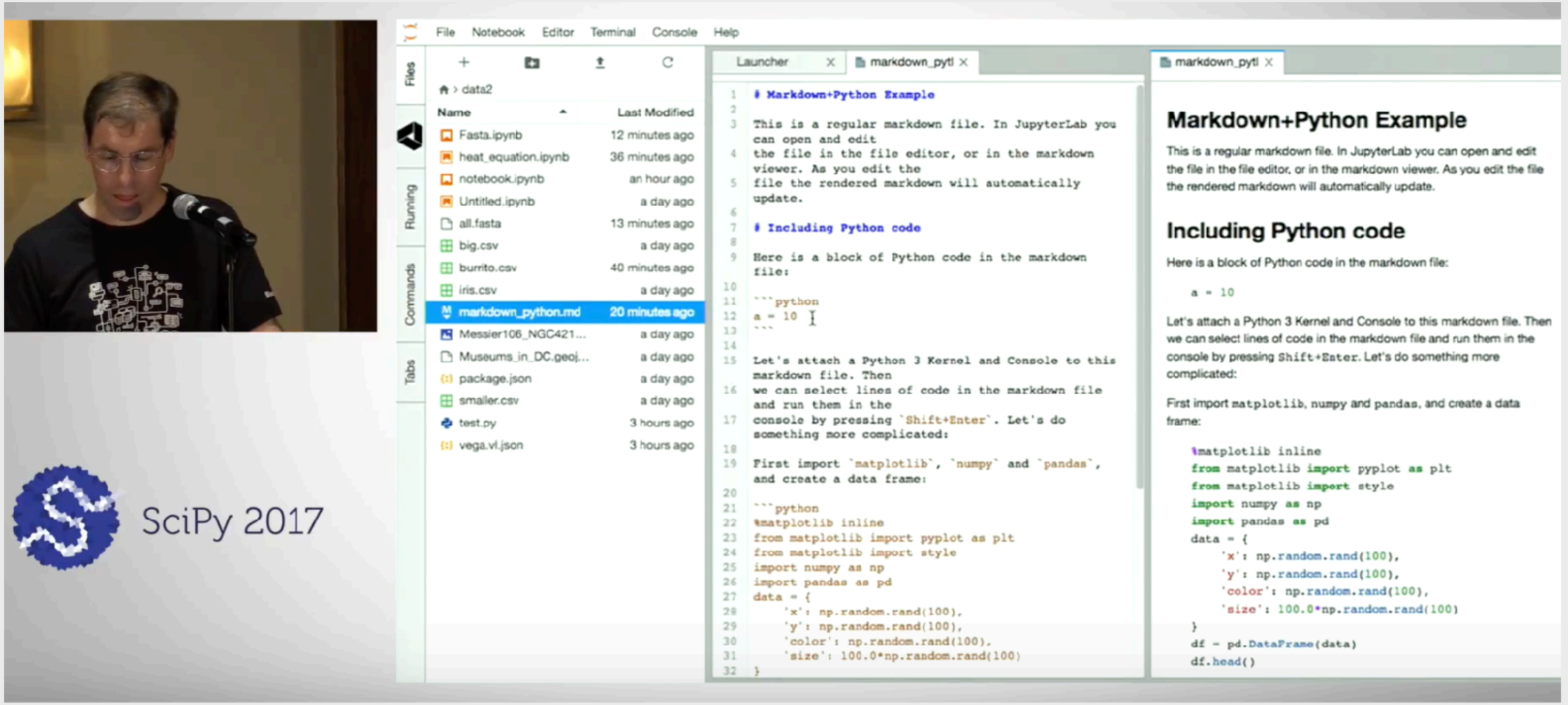
Collaboration between tools



A log in the console of commands executed

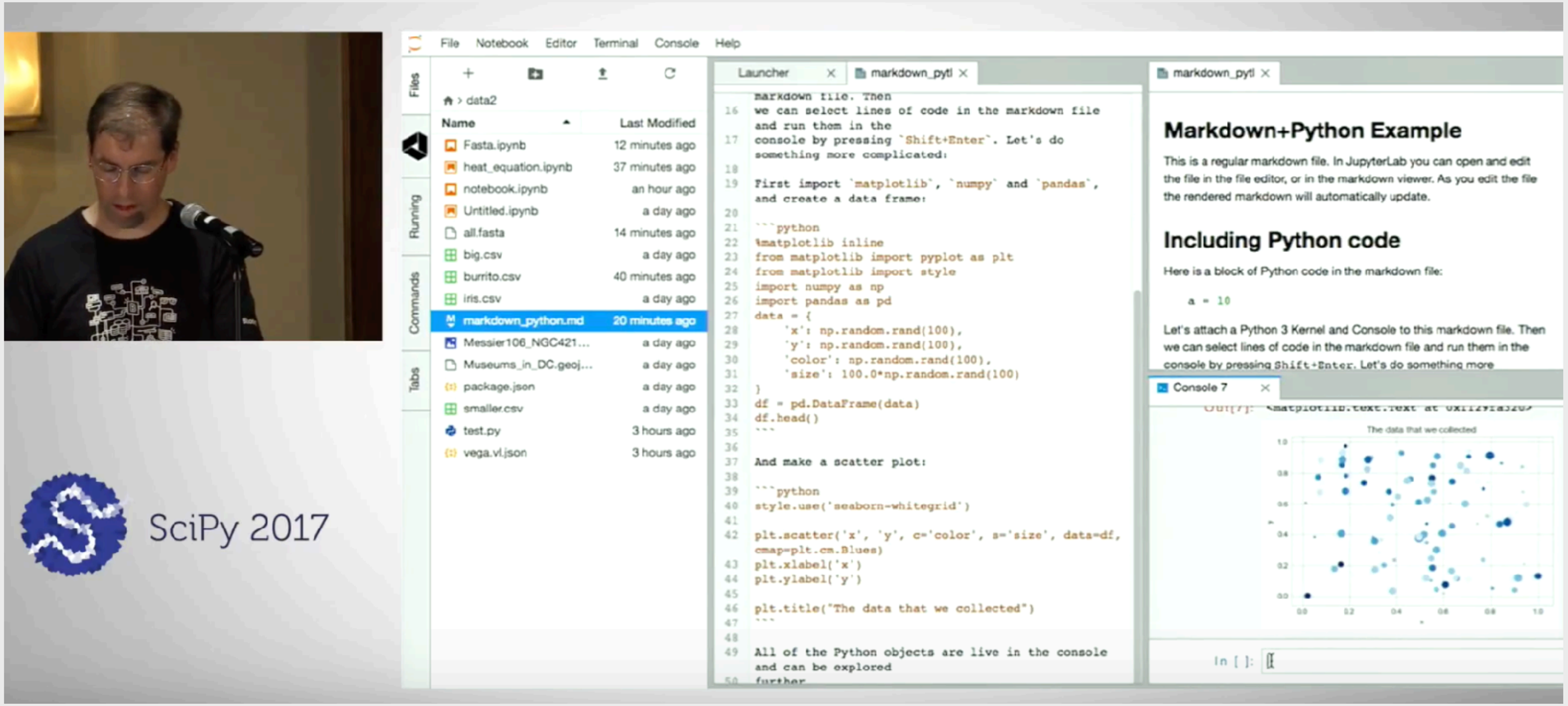
Explore data in console without messing up your notebook

Editors



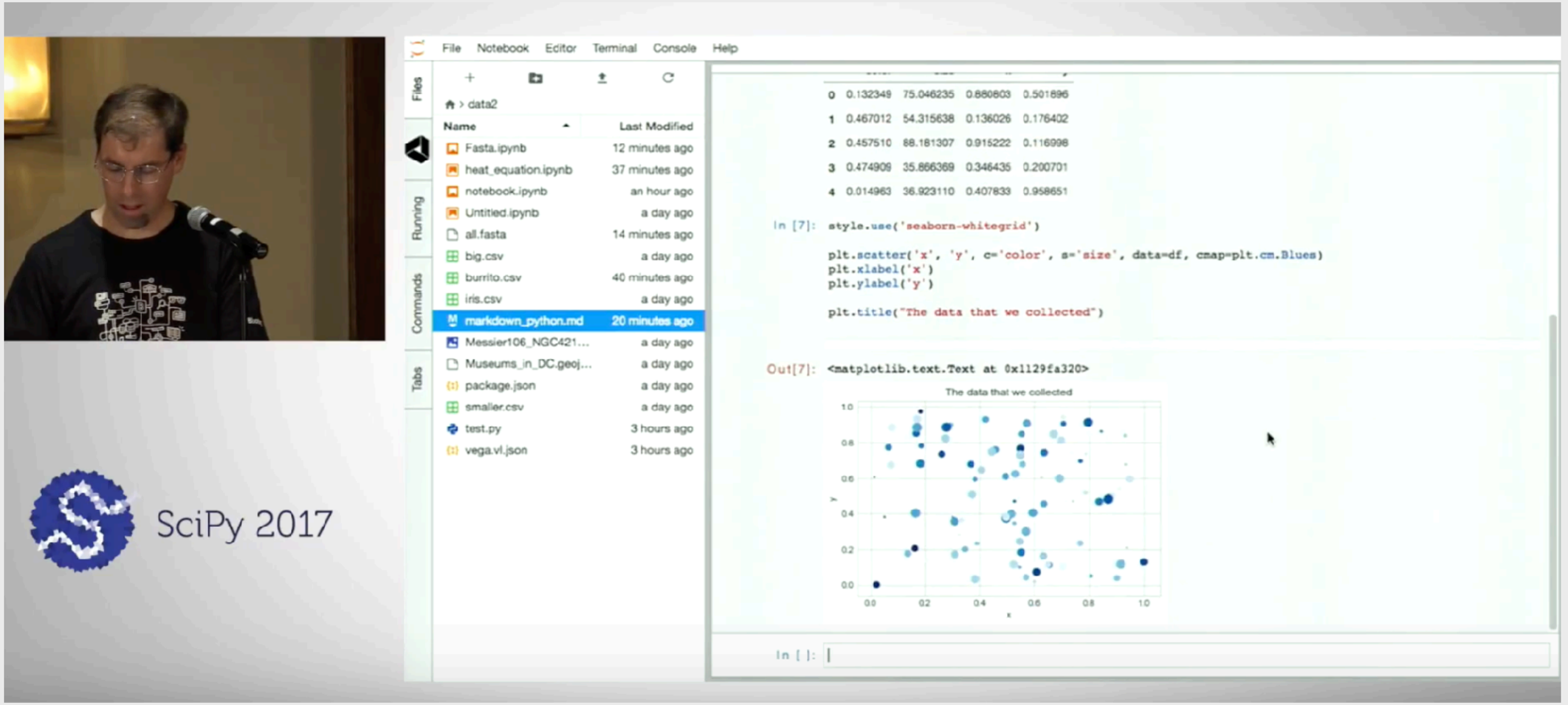
Many different editors; preview markdown

Editor connected to a console



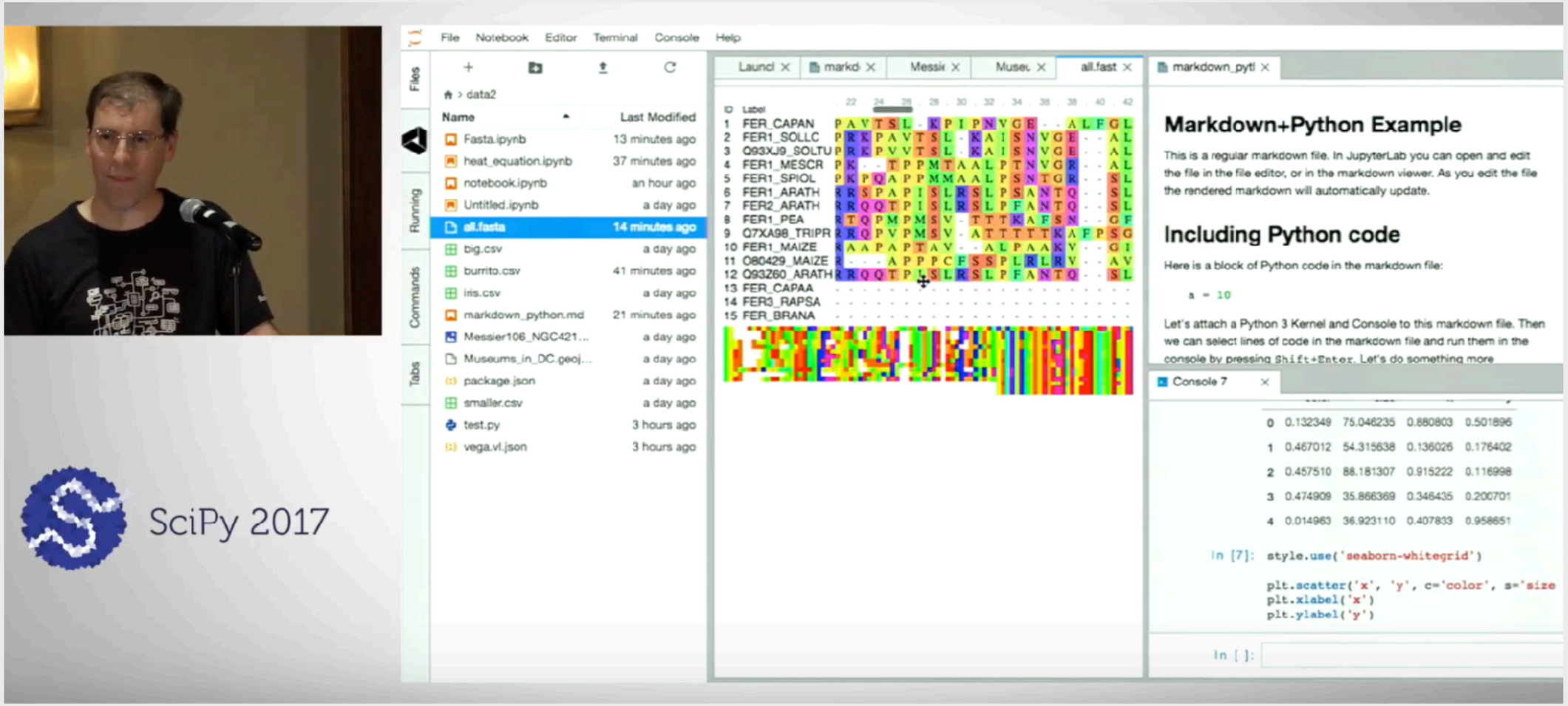
Connect to console and Shift-Enter to run code snippet

Single document mode



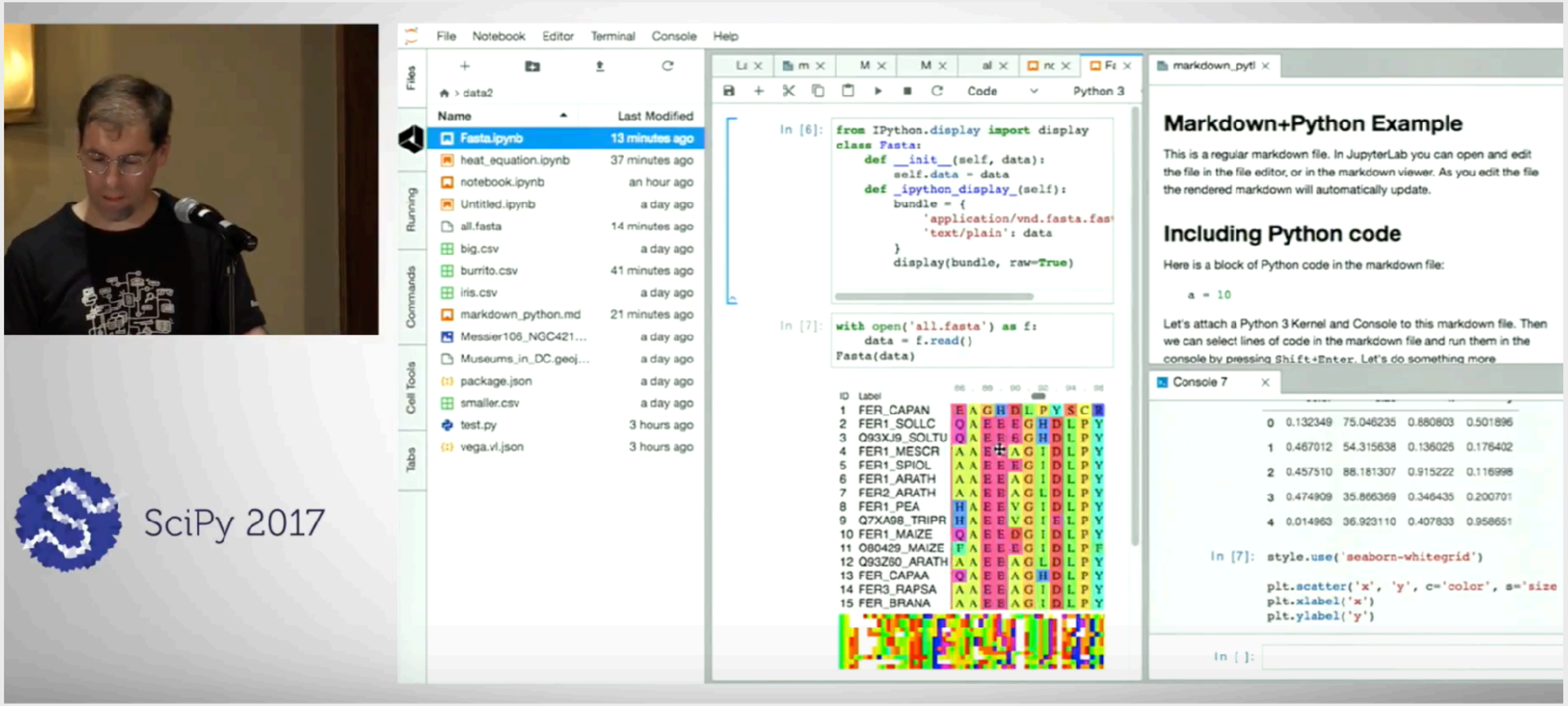
Shift-Command-Enter to enter single document mode. Similar to classic notebook.

Extensible



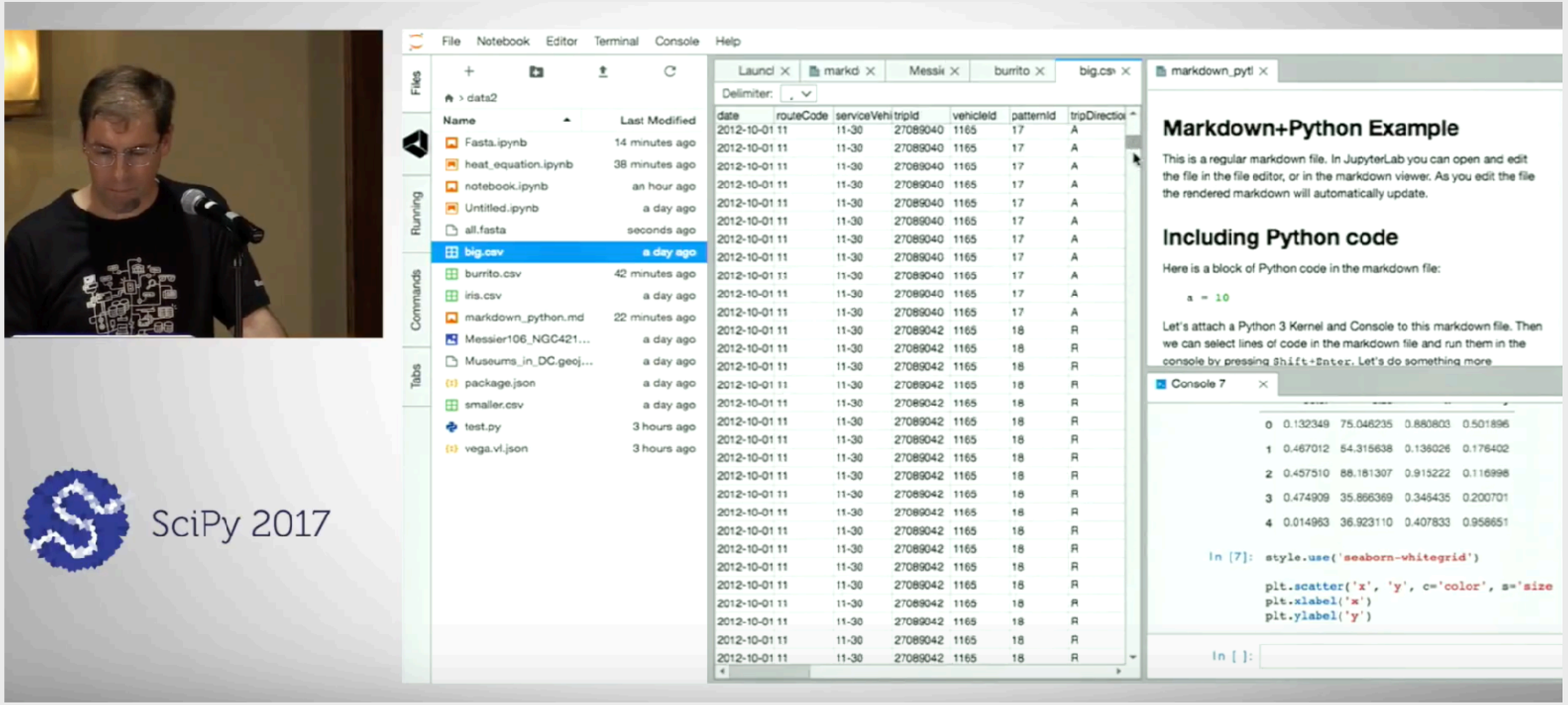
"In one night and a couple of dozen lines of code we wrote a Fasta viewer."

Becomes a notebook extension



With the same code, the Fasta viewer becomes an extension usable in the notebook.

Datasets, grids, and scale



1.2M rows 200Mb csv file. Excel can't open.
A few seconds to load and then "smooth as butter" when scrolling.
Rumor has it that Chris Colbert has a trillion row by column demo too.

Agenda

- Gateways
- Jupyter Notebook
- JupyterHub
- JupyterLab
- Next steps

Call to action

- Join **Jupyter** mailing lists
- Participate in a **sprint**
- Give a **talk** or **write** a post
- Offer a **workshop**
- **Contribute** to a favorite project
- **Share** your trials and successes

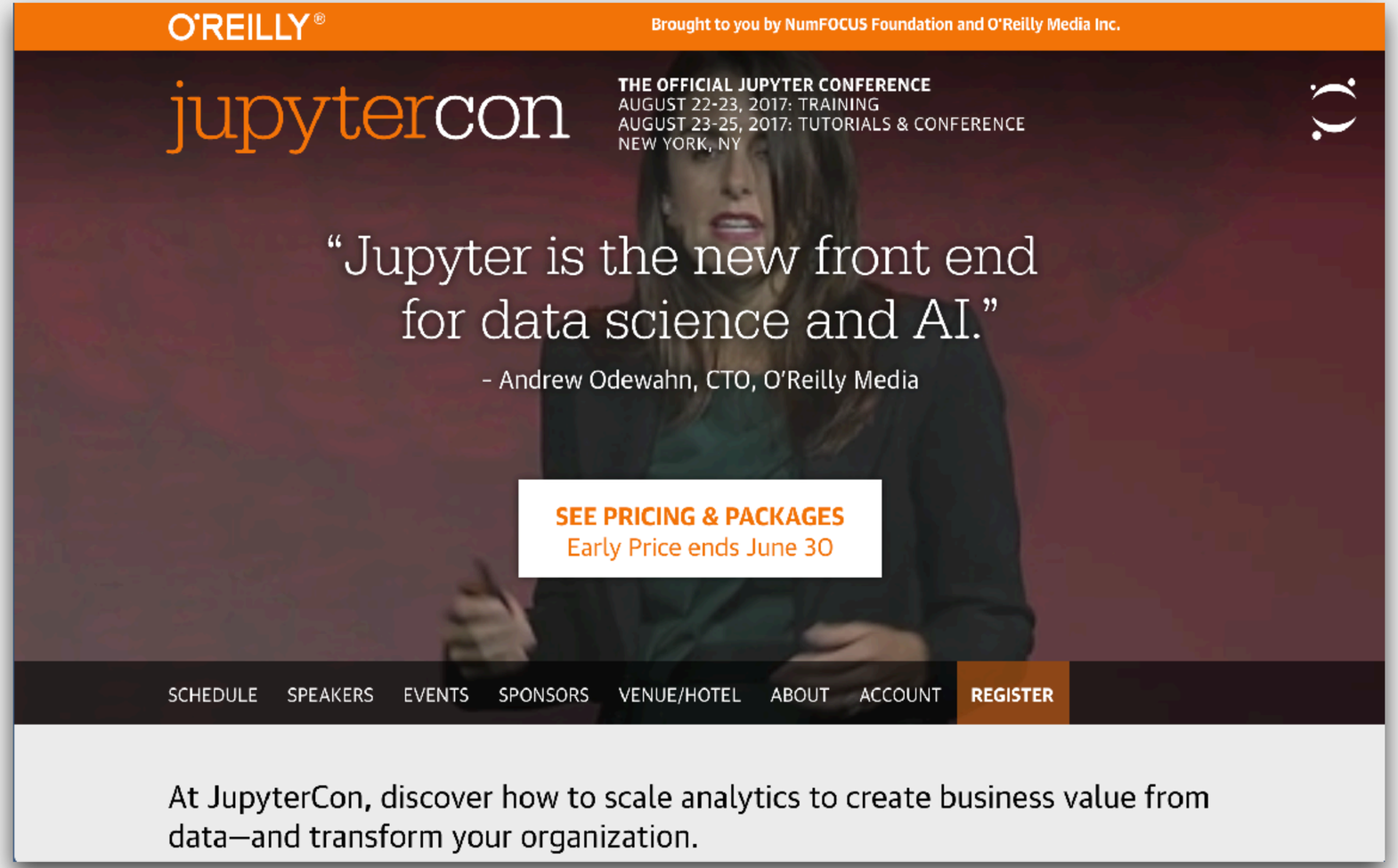
Resources

https://github.com/willingc/2017-science-gateways/blob/master/resources/resources.md

jupyter.org	pyvideo.org
jupyter google groups and Gitter	try.jupyter.org
Trending notebooks on GitHub	nbviewer

JupyterCon, August 2017, NYC

https://conferences.oreilly.com/jupyter/jup-ny



WE'RE OPEN FOR YOU.

Attributions and recognition

A huge thank you to the Project Jupyter team and community.
Your hard work and passion makes this all possible.

GORDON AND BETTY MOORE FOUNDATION

Alfred P. Sloan FOUNDATION

THE LEONA M. AND HARRY B. HELMSLEY CHARITABLE TRUST

- Kristen Thyng
- San Diego Python
- Demba Ba
- Jeremy Freeman, Binder
- Michael Cuthbert, music21
- LIGO
- Andrea Zonca, SDSC, Ilkay Altinas, Software Carpentry
- Photo credits on individual slides

Thank you

Thank you

try.jupyter.org

www.jupyter.org

ipython.org

numfocus.org

GitHub: willingc

@willingcarol

Thank you for participating!

- Please share your feedback through our 30-second evaluation: http://sciencegateways.org/webinareval
- Join us next month (August 9)

  Interactive Best Practices:
  Job Management & Scheduling
  Presented by Miron Livny and Todd Tannenbaum (Condor Project), Mark Miller (CIPRES),Sudhakar Pamidighantam (SEAgrid), and others
- Upcoming opportunities for students/educators: http://sciencegateways.org/engage/student-focused