

# Professor Grumpy's guide to the **conda-jupyter-GitHub** “ecosystem”, eased by our science's maven **Unidata**

Brian Mapes, UM-RSMAS

February 2018



Search it up... (

Guest

# It's a character type, I discover

**Grumpy Professor** is an NPC in Elder Scrolls Online. Non-player Characters are AI that players may interact with during their adventuring in ESO. Some offer quests, others lore insights, some keep Shops and provide vending of items and supplies, and some are simply around for atmosphere building.



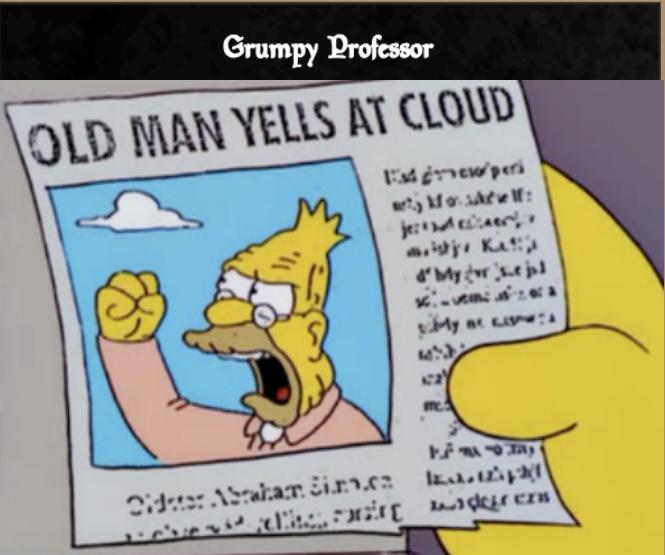
## Grumpy Professor Quest Information

This **ESO character** is related to the following Quests:

- ??
- ??



## Grumpy Professor Notes

 Grumpy Professor	
Race	??
Gender	Male
Location	• Auridon
Related Quests	??
Level / hp	50

# Computer needs for science career

- Computing and graphics production
  - Keeping track of versions
- Backup against loss
- Reproducibility
  - by self, later
  - by others, elsewhere
- Collaboration
  - with group (& just double checking student work...)
  - across time and space
    - rarely simultaneous collisions
- Communication
  - teaching of students
  - sharing of research results (teaching of peers)

# Computer needs for science career

- Computing and graphics production
    - Keeping track of versions
  - Backup against loss
  - Reproducibility
    - by self, later
    - by others, elsewhere
  - Collaboration
    - with group (& just double checking student work...)
    - across time and space
      - rarely simultaneous collisions
  - Communication
    - teaching of students
    - sharing of research results (teaching of peers)
- Advisor's favorite licensed language naming hygiene (or not) Dropbox/Box, OneDrive, etc.
- naming hygiene (or not), file/email time stamps
- cross-checks at figure level mostly RTFC: code *is* documentation
- emailing codes
- paste figs into PowerPoint/Keynote, Word/ LaTeX, etc.

# conda-juPy-GitHub “ecosystem”

- Computing and graphics production – Keeping track of versions *literate programming* git-- requires commit msgs.
- Backup against loss [github.com](https://github.com)
- Reproducibility
  - by self, later
  - by others, elsewhere[conda envs. & github.com](#)
- Collaboration
  - with group (& juust double checking student work...)
  - across time and space
    - rarely simultaneous collisions[github.com](https://github.com)
- Communication
  - teaching of students
  - sharing of research results (teaching of peers)*jupyter,*  
[github.com](https://github.com)  
+ all your usual “documents” apparatus... as needed... for now...

# You lost me at “free”

- Will it last?
  - vanity project of aging 20<sup>th</sup> century cyber-socialists, supported by charity or Silicon Valley largesse?
- Do I have to memorize a bunch of commands that are a mashup of some guy’s favorite fantasy heroine and his seventh cat’s name?
- When inscrutable commands fail with gibberish errors, will the authors of every part say it must be just *my* problem, involving the other parts?
  - or maybe I’ll find 23 ad-filled question sites with the same question & 59 inarticulate half-answers?



# Let's tackle that first question, Grumpy

- Will it last?
  - vanity project of aging 20<sup>th</sup> century cyber-utopians, supported by charity or Silicon Valley largesse?
- I think it will last.
- GitHub.com is *free to open uses*; pay for privacy.
- Anaconda.com distributes juPy for *free to all*, makes their money on building open-source “solutions”
- Good relationship between public & private
  - science & education get software-quality tools
  - software industry gets people trained on those tools
  - breadth (openness) makes everyone’s stuff robust



# And commerce is better?

- Finally updated Microsoft Office 2016 this week
- Massive! 8 GB! Well, it must be good.
- Used it to write this talk. Double check before presentation: **\$@^\*# Microsoft !!!**

 The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears:

The screenshot shows a Microsoft TechCenter forum page. At the top, there's a navigation bar with the Microsoft logo, "TechCenter", a search bar ("Search Office with Bing"), and user information ("United States (English) Sign out miamiprof"). Below the navigation, there are links for "Home", "2013", "2010", "Other Versions", "Library", "Forums" (which is highlighted in red), and "Gallery". There are three main buttons: "Ask a question" (red), "Search related threads" (light blue), and "Search forum questions" (grey). On the left, there's a "Quick access" dropdown menu showing a profile picture for "scgsg" (5 points, joined Apr 2011), a link to "scgsg's threads", and a "Show activity" button. The main content area displays a post titled "Powerpoint 2016 - the picture cannot be displayed" with a red feed icon. Below the title, it says "Microsoft Office > Office 2016 and Office 365 ProPlus – IT Pro Discussions". The post has a "Question" button and an "Alert me" button. A detailed description of the issue follows: "The powerpoint files are ppt files and I've tested them with a Win7 machine with Office 2010 on it. The images are displayed fine but on computers with powerpoint 2016 get empty boxes with a red x with the error 'the picture cannot be displayed'. I've also found that if I open the file on a Win7 machine with Office 2010, and save the powerpoint as pptx, files can be opened in 2016".

# For the usability questions, we need a **maven**

- Will it last?
  - vanity project of aging 20<sup>th</sup> century cyber-utopians, supported by charity or Silicon Valley largesse?
- Do I have to memorize a bunch of commands that are a mashup of some dude's favorite fantasy heroine and his seventh cat's name?
- When inscrutable commands fail with gibberish errors, will the authors of every part say it must be just *my* problem, involving the other parts?
  - or maybe I'll find 23 ad-filled question sites with the same question & 59 inarticulate half-answers?



# ma·ven

/'māvən/ ⓘ

**noun** NORTH AMERICAN *informal*  
noun: **maven**; plural noun: **mavens**

an expert or connoisseur.  
"fashion mavens"

## Origin

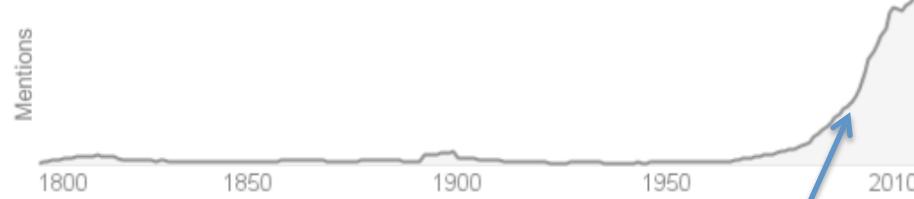
YIDDISH

→ maven  
1960s

1960s: Yiddish.

Translate maven to

Use over time for: maven



Malcolm Gladwell popularised the term more widely in his book *The Tipping Point* (Little Brown, 2000). Gladwell described those who are intense gatherers of information and also impressions, and so are often the first to become aware of new or nascent trends. The word has since become widely used in particular

## History [ edit ]

The word itself is a borrowing from the [Yiddish](#) *meyvn* 'an expert, connoisseur', derived from the [Hebrew](#) מִבֵּן *mēvīn* 'person with understanding, teacher', a participle of the verb הָבֵין *hēvīn* 'to understand'<sup>[1]</sup>, from the [West Semitic root](#) *byn* 'to be separate, distinguish'<sup>[2]</sup>.

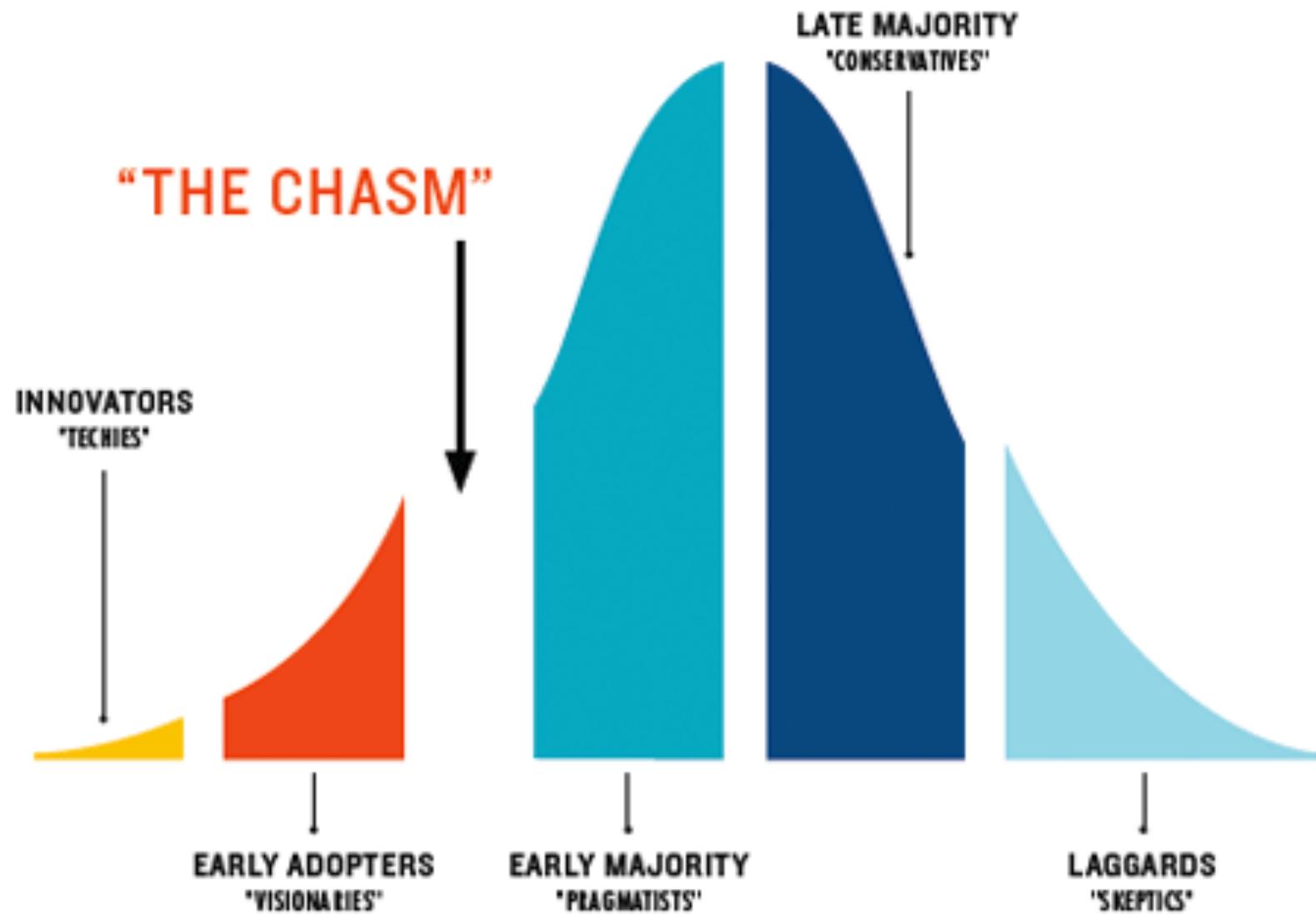
It was first recorded (spelled *mayvin*) in [English](#) in 1950 (in the *Jewish Standard* of [Toronto, Ontario, Canada](#)) and popularized in the United States in the 1960s by a series of commercials created

# My maven: Unidata

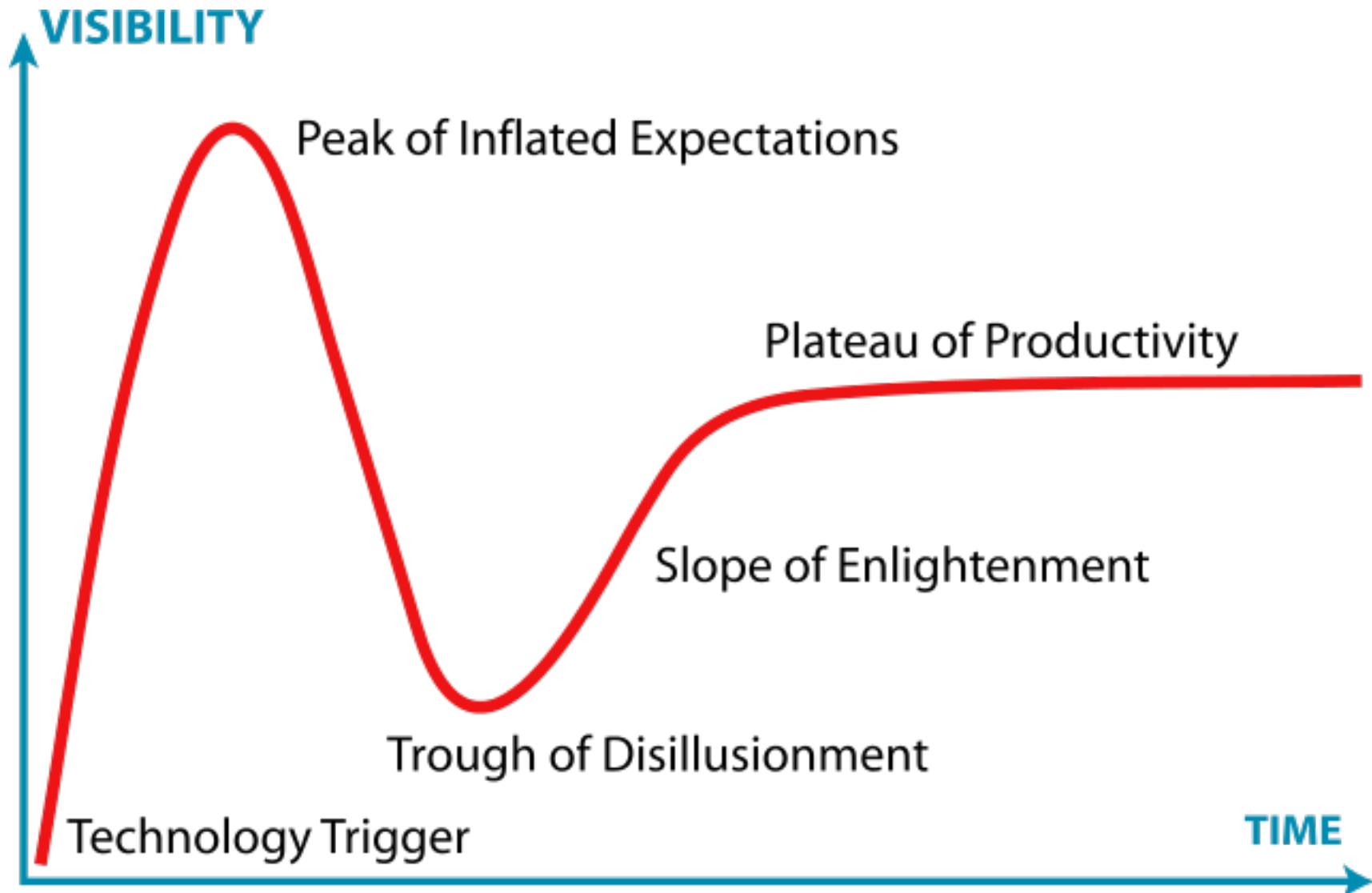
The screenshot shows the homepage of the Unidata website at https://www.unidata.ucar.edu. The page features a dark blue header with the UCAR Community Programs logo and the Unidata logo. The Unidata logo includes the text "Data Services and Tools for Geoscience". Below the header, there's a banner with text about regional workshops and a video thumbnail showing a person speaking. To the right, there's a "Welcome to Unidata" section with a "Learn more about" button. The main menu at the top includes links for Data, Software, Downloads, Support, Community, Projects, News, Events, About Us, Login, and Register. The address bar shows the secure connection and the URL.

- But RSMAS has local ones too...
  - Brazilians! Leo, Tiago, & earlier
    - » before I (and maybe the technology) were quite ready...
  - Europeans! Milan, Ray Bell, ...
- It is *much* bigger than just our science...

I think **conda**-**juPy**-**GitHub** has crossed **THE CHASM**  
[https://en.wikipedia.org/wiki/Technology\\_adoption\\_life\\_cycle](https://en.wikipedia.org/wiki/Technology_adoption_life_cycle)

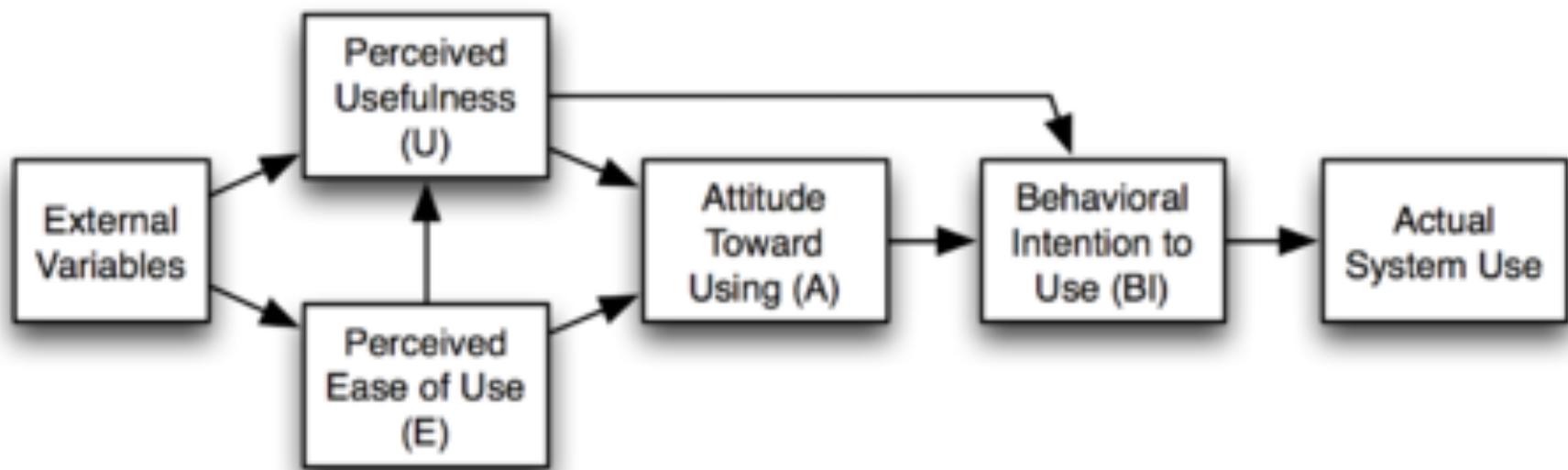


And levelled out into **Productivity**  
[https://en.wikipedia.org/wiki/Hype\\_cycle](https://en.wikipedia.org/wiki/Hype_cycle)



# So let's roll on...

[https://en.wikipedia.org/wiki/Technology\\_acceptance\\_model](https://en.wikipedia.org/wiki/Technology_acceptance_model)

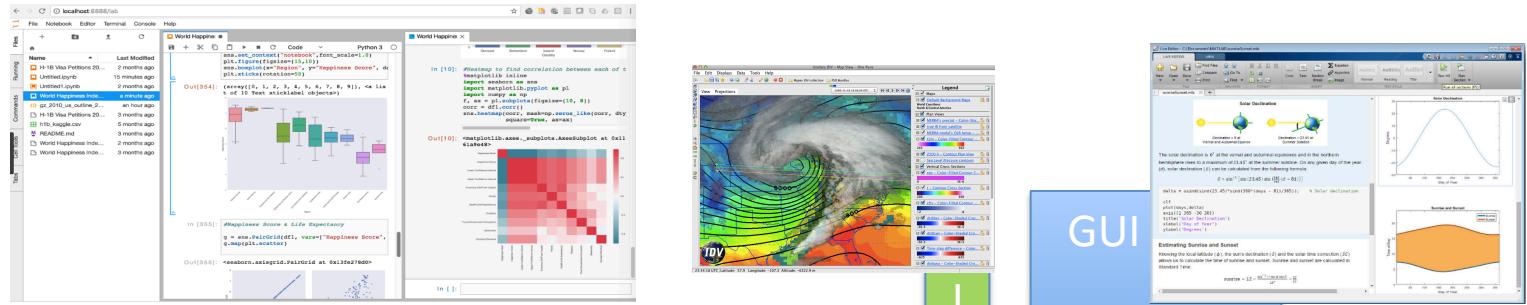


# Outline

- Before bushwhacking, survey the landscape
  - The IT Stack: hardware, platforms, API ‘levelers’, and user interfaces for scientists
  - Python world, the realm of science, and where a web search tends to drop you
- *Why all of conda-jupyter-GitHub is worth it*
- Where to find **conda-jupyter-GitHub onramp**
- Showcase (finish in a notebook)

# Leveling a rugged IT landscape so science can do science

Graphical  
Interfaces



Interpreted  
(scripting)  
compilers



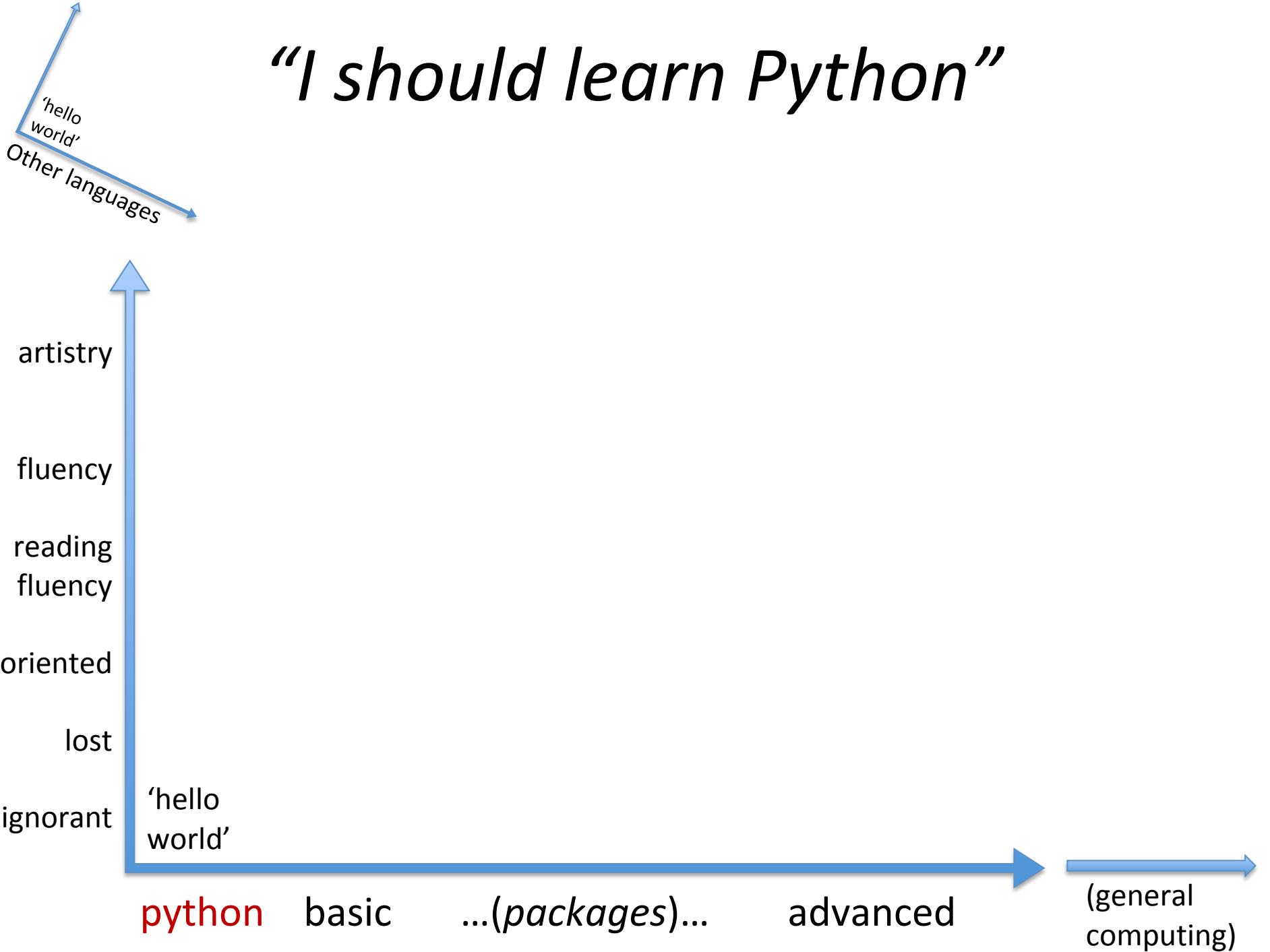
Browser wars  
From Wikipedia,  
JavaScript.c

OS

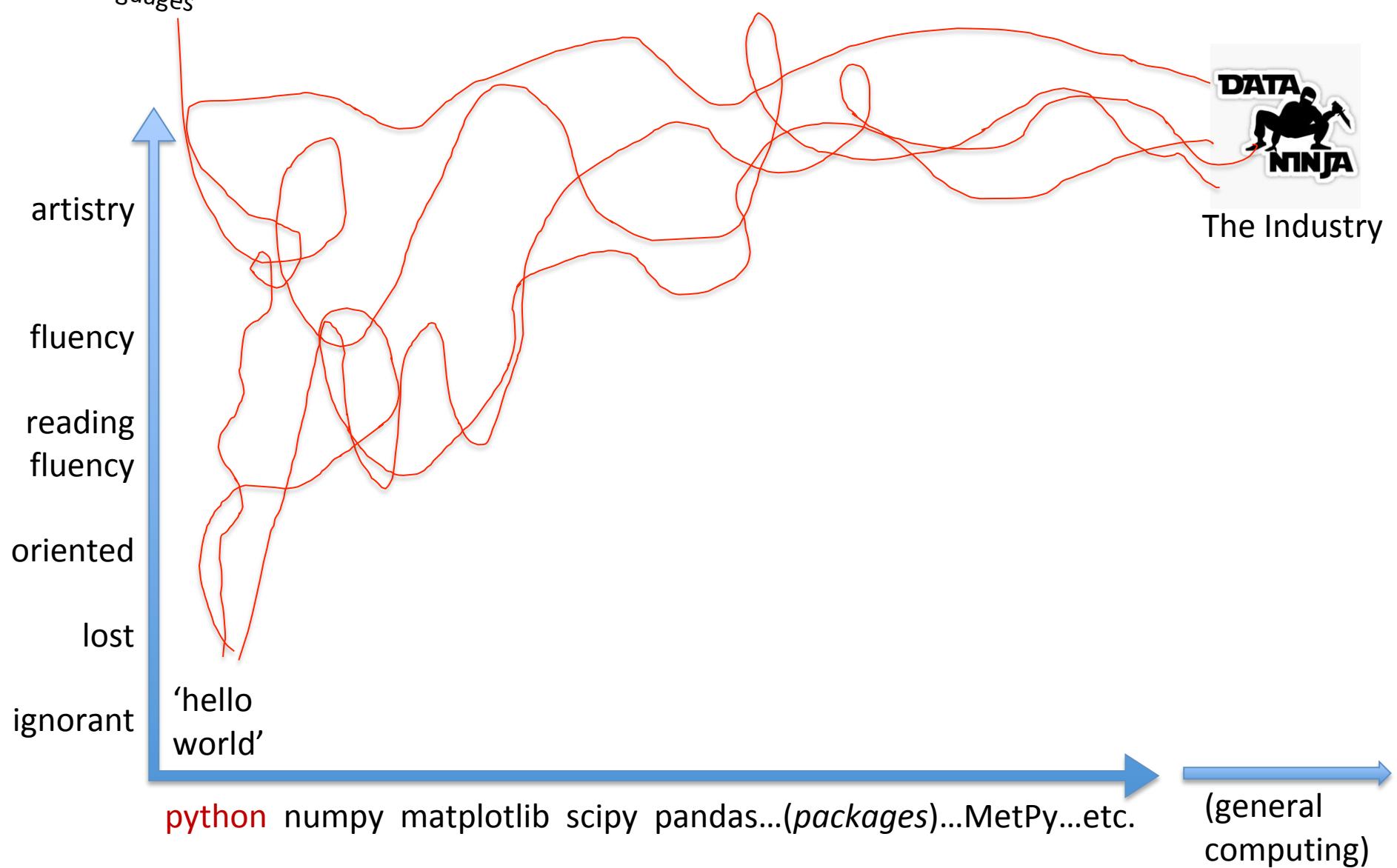


Hardware

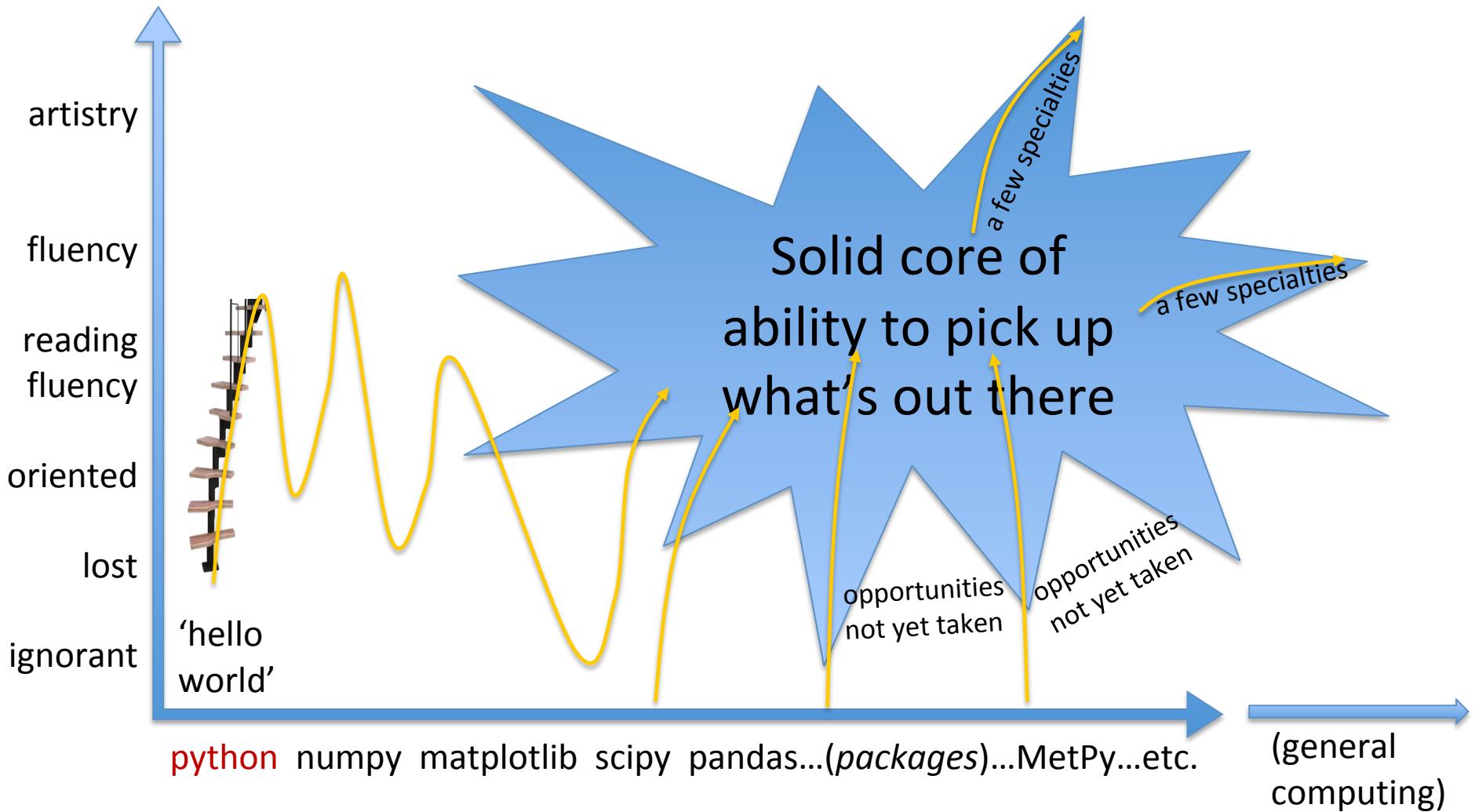




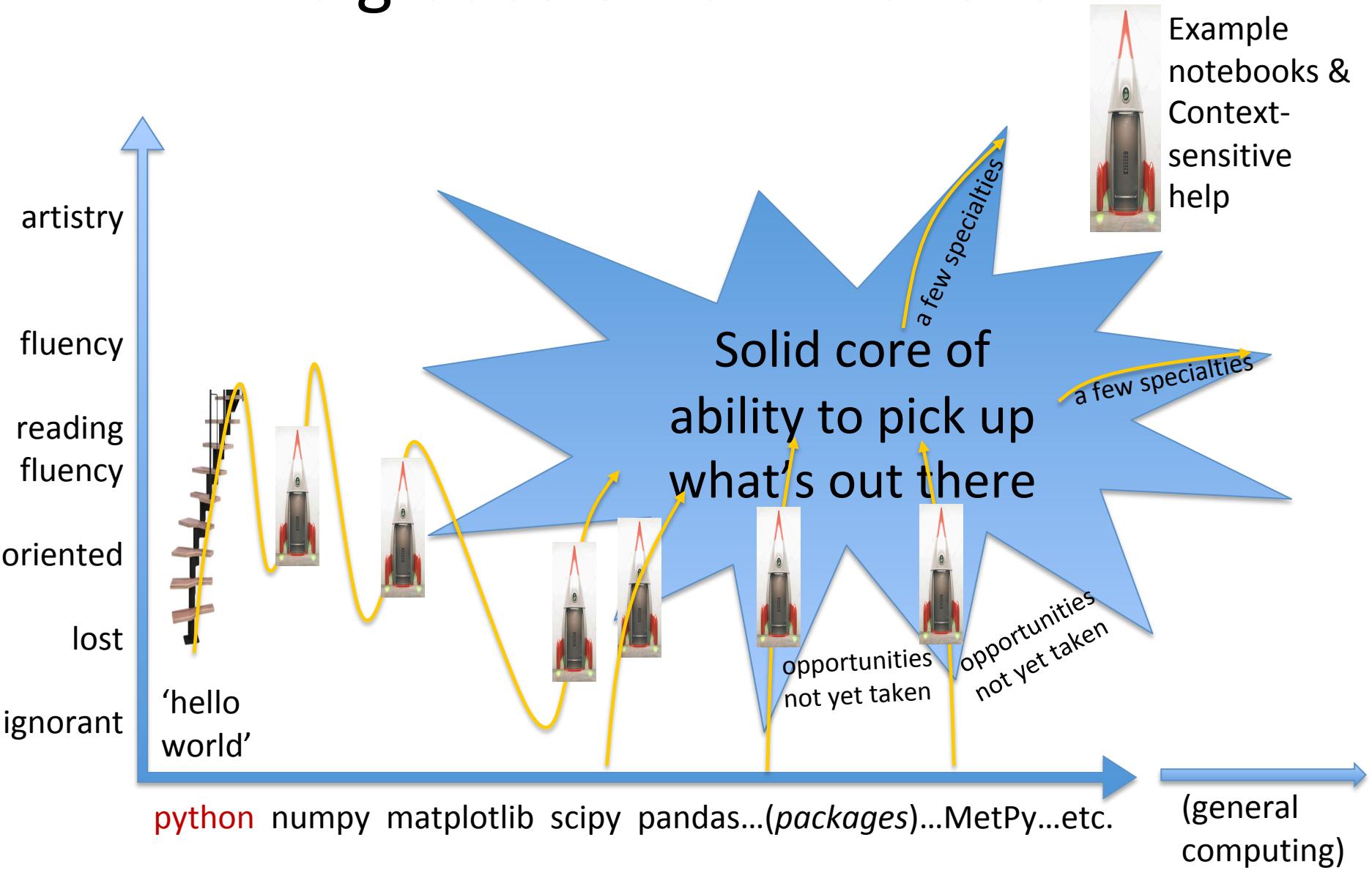
# “Data science” path: *populous*, so *it’s where web search tends to drop you*



# Scientist path: a career-making skill set



# After that first staircase, can get big boosts from mavens!



## Advanced Pythonic Data Analysis

Unidata Python Workshop



### Overview:

- Teaching: 45 minutes
- Exercises: 45 minutes

### Questions

1. How can we improve upon the versatility of the plotter developed in the basic time series notebook?
2. How can we iterate over all data file in a directory?
3. How can data processing functions be applied on a variable-by-variable basis?

### Objectives

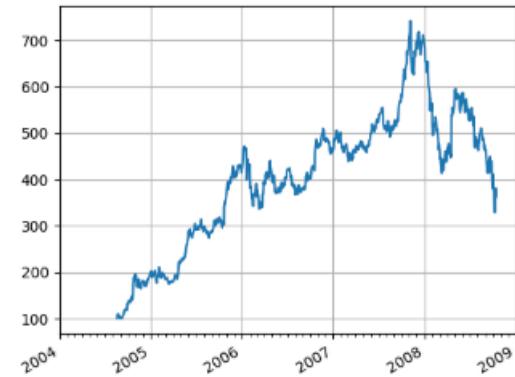
1. [From Time Series Plotting Episode](#)
2. [Dictionaries of Parameters](#)
3. [Multiple Dictionaries](#)
4. [Function Application](#)
5. [Glob and Multiple Files](#)

### From Time Series Plotting Episode

Here's the basic set of imports and data reading functionality that we established in the [Basic Time Series Plotting](#) notebook.

In [ ]: `from datetime import datetime  
import pandas as pd`

Elegant style examples:  
header w/ images,  
html sections,...



Timed lessons!  
Links between lessons

# Unidata's examples

**EXERCISE:** As a final exercise, use a dictionary to allow all of the plots to share

In [ ]: *# Your code goes here*

In [ ]: *# %load solutions/pythonic\_bounds.py*

[Top](#)

---

Exercises, with solution-reveal:

# Mavens make all the difference

- really, *community*, once you are boosted...

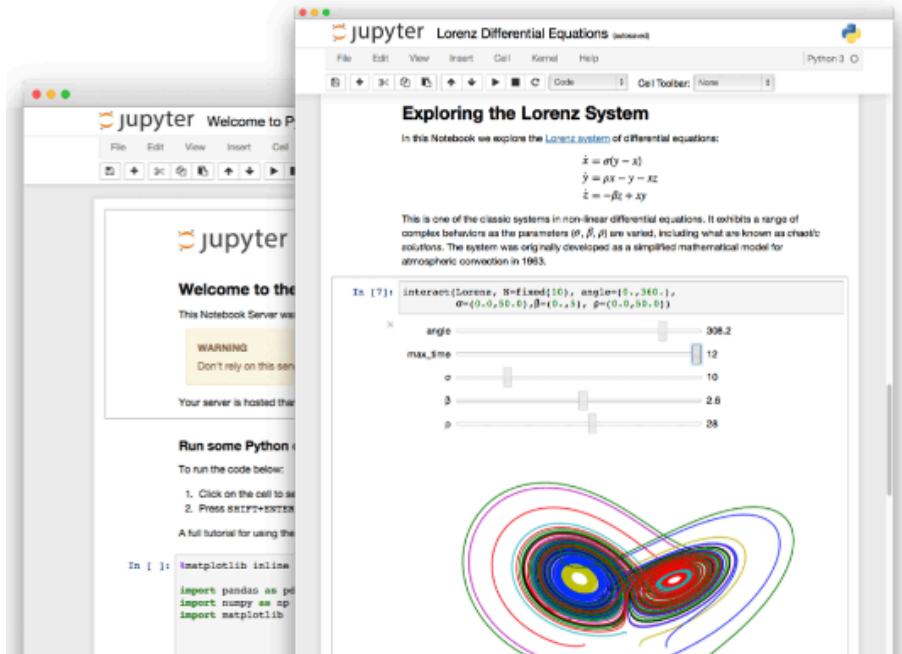
# Outline

- Appreciating the landscape
  - The IT Stack: hardware, platforms, API ‘levelers’, and user interfaces for scientists
  - Python world, the realm of science, and where a web search tends to drop you
- *Why all of conda-juPy-GitHub is worth it*
- Where to find **conda-juPy-GitHub onramp**
- Showcase (finish in a notebook)

# conda -juPy-GitHub

- “Conda is an open source package management system and environment management system...created for Python programs, but ... for any language.”
  - » anaconda.com is the company behind it
- **Free software distributor!**
  - `jupyter` and `Python` + `packages` all come from here (`pip` too)
- “Environment management system” is about as fun as it sounds... *Professor Grumpy's biggest grrr*
- But,
  - Package ***combinations*** are the key to Python’s modular power & the collaborative vision.
  - Keeping track of these is the key to ***reproducibility***
    - » and...my unfavorite unintuitive command will be renamed soon, I hear...

conda-jupyter-GitHub



## The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

[Try it in your browser](#)

[Install the Notebook](#)

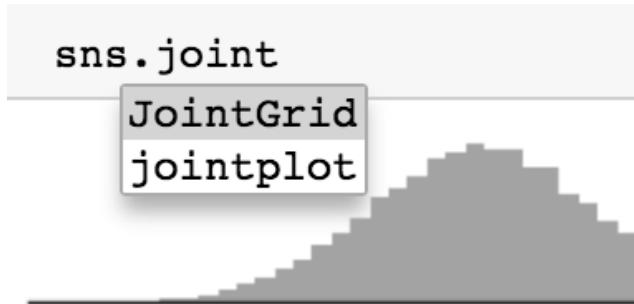
- **jupyter notebook** process manages OS-level resources
  - streams its I/O to port 8888, in browser-ese
- **Your browser** manages the human- screen- keyboard- mouse interactions
  - *stable, free infrastructure*



- Formerly “iPython notebook”
  - file suffix remains .ipynb
    - & many services start with nb (nbconvert, etc.)
- Jupyter generalization: 2014
  - Ju refers to Julia (a python-like language -- faster)
  - the R is for that statistics-oriented language (formerly S, S+)
- Now very general: has kernels (interpreters) for dozens of languages (e.g. octave, & matlab’s python engine)

# conda-jupyter-GitHub

- A light Development Environment, with inline help
  - TAB shows *methods and attributes* available for any object



- SHIFT+TAB shows the *docstring* for any object

```
Signature: sns.jointplot(x, y, data=None, kind='scatter', stat_func=<function pearsonr at 0x11b12a9d8>, color=None, size=6, ratio=5, space=0.2, dropna=True, xlim=None, ylim=None, joint_kws=None, marginal_kws=None, annot_kws=None, **kwargs)
Docstring:
Draw a plot of two variables with bivariate and univariate graphs.

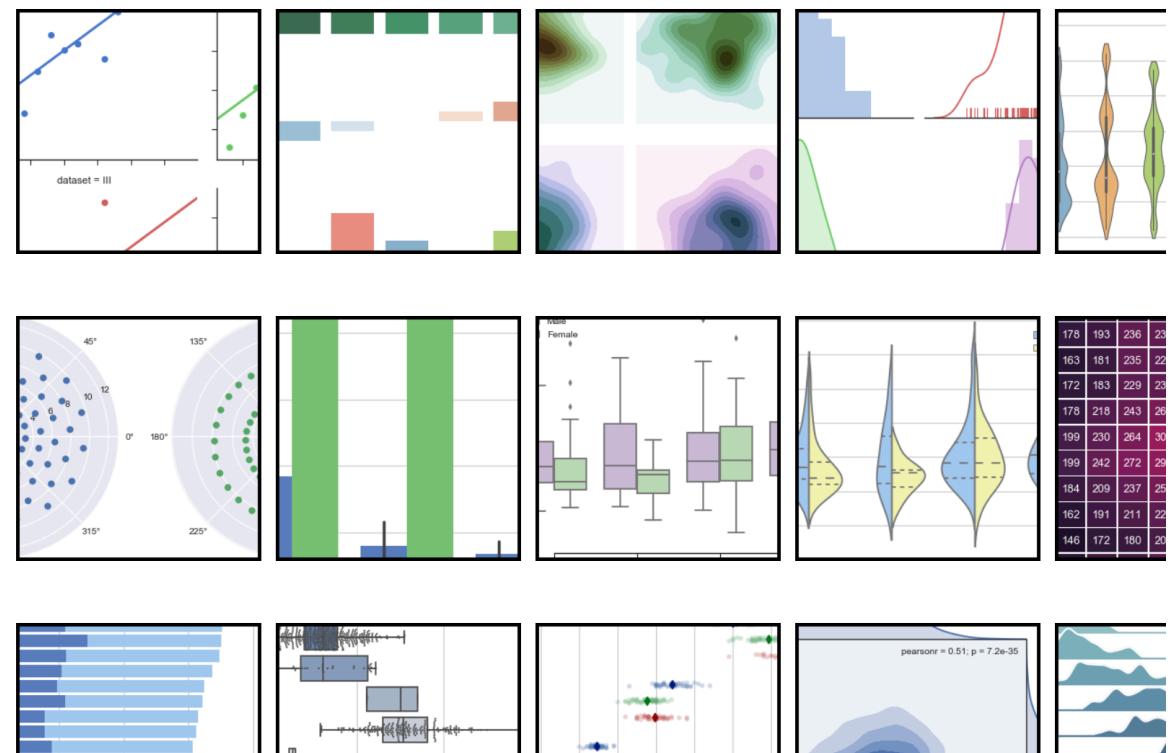
This function provides a convenient interface to the :class:`JointGrid` class, with several canned plot kinds. This is intended to be a fairly lightweight wrapper; if you need more flexibility, you should use :class:`JointGrid` directly.
```

# conda-ju Py-GitHub

- “galleries” let you work backward from a result-shaped container

seaborn 0.8.1 Gallery Tutorial API Site ▾ Page ▾

## Example gallery



# conda-juPy-GitHub

- “GitHub is a web-based hosting service for ... git.”
  - “Git is a version control system ... created by Linus Torvalds in 2005 for development of the Linux kernel.”
- As of April 2017, GitHub reports having almost 20 million users and 57 million repositories, making it the largest host of source code in the world. (Wikipedia)
- GitHub has a mascot, an "Octocat" called Mona, a cat with five tentacles and a human-like face.



# conda-juPy-GitHub

- **Big point for this: GitHub.com renders .ipynb files as static viewable pages**
  - Better yet: install nbviewer.jupyter.org browser extension
- Repositories of .ipynb files are an *amazing resource!*
  - <https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>
  - With context-sensitive help, quickly learn to adjust & adapt anything you can find in an example
  - With Python's module structure, mix-n-match
    - initial namespace collisions? manage them in `conda`
- **#NeverAgainStartFromScratch !**



# Kids have it too easy



- **#NeverAgainStartFromScratch ??**
- mumble **BLACK BOX** grumble harumph
  - » not my brand of grumpy, so I can't express it well

# conda-juPy-GitHub: tiers of commitment

## 1. Scavenger: know how to operate the Web site

MPOcanes / MPO624-2018

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master MPO624-2018 / notebooks / Jupyter Notebook Users Manual.ipynb

brianmapes Add some notebooks

1 contributor

2409 lines (2408 sloc) | 85.2 KB

Raw Blame History

Learn to see this!  
(one file)

This page describes the functionality of the [Jupyter](#) electronic document system. Jupyter documents are called "notebooks" and can be seen as many things at once. For example, notebooks allow:

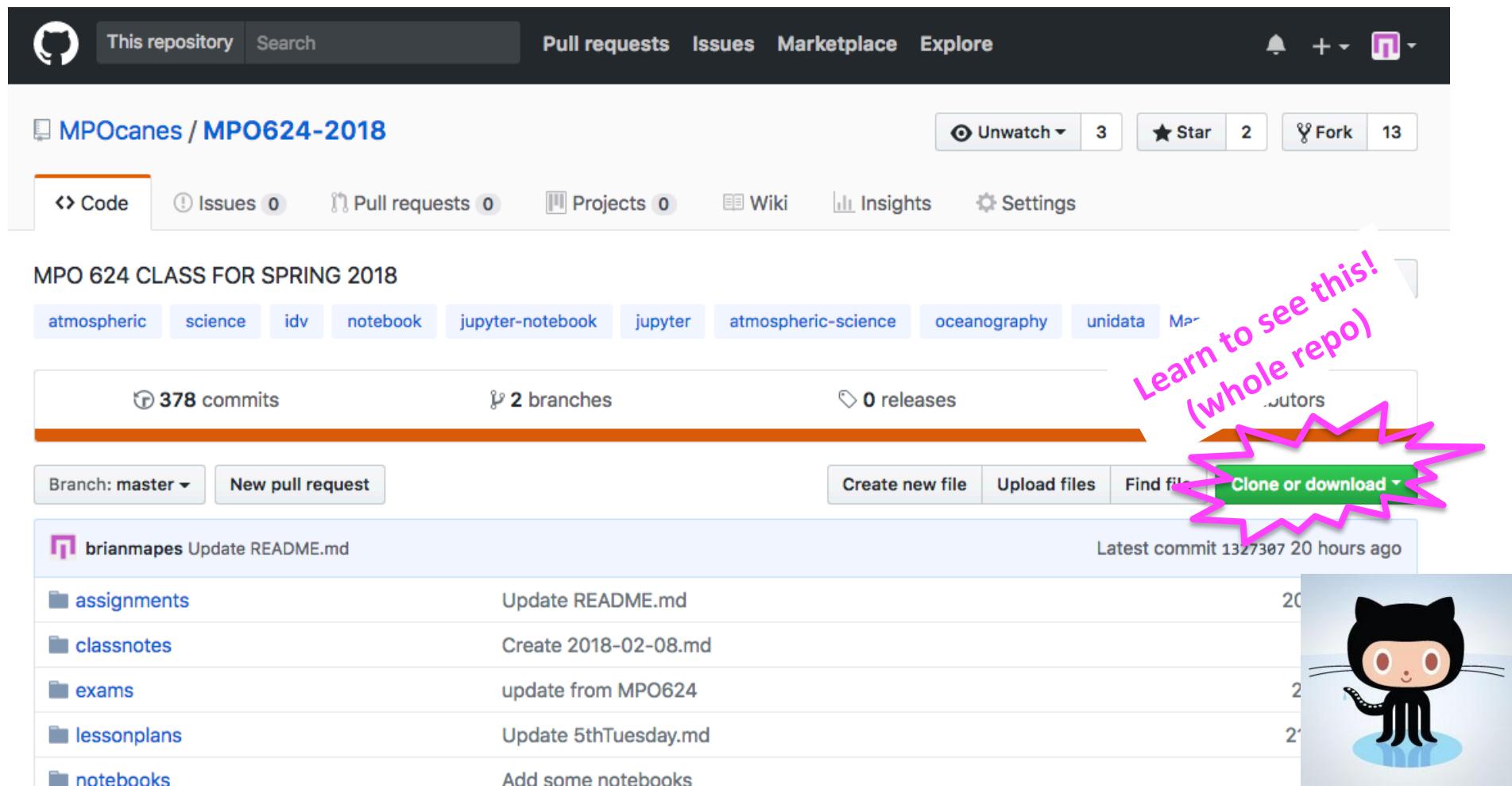
- creation in a **standard web browser**
- direct **sharing**
- using **text with styles** (such as italics and titles) to be explicitly marked using a [wikitext language](#)
- easy creation and display of beautiful **equations**
- creation and execution of interactive embedded **computer programs**
- easy creation and display of **interactive visualizations**

Jupyter notebooks (previously called "IPython notebooks") are thus interesting and useful to different groups of people:



# conda-jupyter-GitHub: tiers of commitment

## 1. Scavenger: know how to operate the Web site



MPOcanes / MPO624-2018

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

MPO 624 CLASS FOR SPRING 2018

atmospheric science idv notebook jupyter-notebook jupyter atmospheric-science oceanography unidata Map

378 commits 2 branches 0 releases

Branch: master New pull request Create new file Upload files Find file Clone or download

brianmapes Update README.md Latest commit 1327307 20 hours ago

assignments Update README.md  
classnotes Create 2018-02-08.md  
exams update from MPO624  
lessonplans Update 5thTuesday.md  
notebooks Add some notebooks

Learn to see this!  
(whole repo)



# conda-jupyter-GitHub: tiers of commitment

## 1. Scavenger: know how to operate the Web site

The screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The page includes a navigation bar with links to 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the navigation bar, there's a header with a bell icon, a plus sign, and a profile picture. A star icon indicates 2 stars, and a fork icon indicates 13 forks. Handwritten pink annotations include 'popular?' next to the star and fork counts. Another annotation says 'Don't miss these tags, to discover like-minded repositories' pointing to the tags section. The repository details show 2 branches, 0 releases, and 13 contributors. At the bottom, there's a commit history with a green 'Clone or download' button. Handwritten pink annotations ask 'is project dead or alive?' and point to the commit times.

MPOcanes / MPO624-2018

This repository Search

Pull requests Issues Marketplace Explore

Unwatch 3 Star 2 Fork 13

Code Issues Requests 0 Projects 0 Wiki Insights Settings

MPO 624-2018

jupyter-notebook atmospheric-science oceanography griddata Manage topics

Commits 2 branches 0 releases 13 contributors

Branch: master New pull request

Create new file Upload files Find file Clone or download

brianmapes Update README.md

assignments Update README.md

classnotes Create 2018-02-08.md

exams update from MPO624

lessonplans Update 5thTuesday.md

notebooks Add some notebooks

Latest commit 1327307 20 hours ago

20 hours ago 3 days ago 25 days ago 21 hours ago 6 days ago

is project dead or alive?

conda-jupyter GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world

The screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The page includes navigation links like 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the header are sections for 'Code', 'Issues (0)', 'Pull requests (0)', 'Projects (0)', 'Wiki', 'Insights', and 'Settings'. A dropdown menu is open at the top right, showing options: 'New repository', 'Import repository', 'New gist', 'New organization', 'This repository', and 'New issue'. A pink hand-drawn style arrow points from the top right towards the '+' icon in the header. The main content area displays the README.md file, which starts with 'MPO 624 CLASS FOR SPRING 2018' and 'Welcome!', followed by the text 'The course is all here. Fork it so you can contribute back!'. A GitHub logo icon is also present.

# conda-jupyter-GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world

MPOcanes / MPO624-2018

Code Issues 0 Pull requests 0 Projects 0

Branch: master MPO624-2018 / README.md

brianmapes Update README.md

1 contributor

22 lines (14 sloc) | 1.29 KB

.md is  
“markdown”  
easy to edit,  
nice looking formatting!

Unwatch 3 Star 2 Fork 13

Raw Blame History

MPO 624 CLASS FOR SPRING 2018

Welcome!

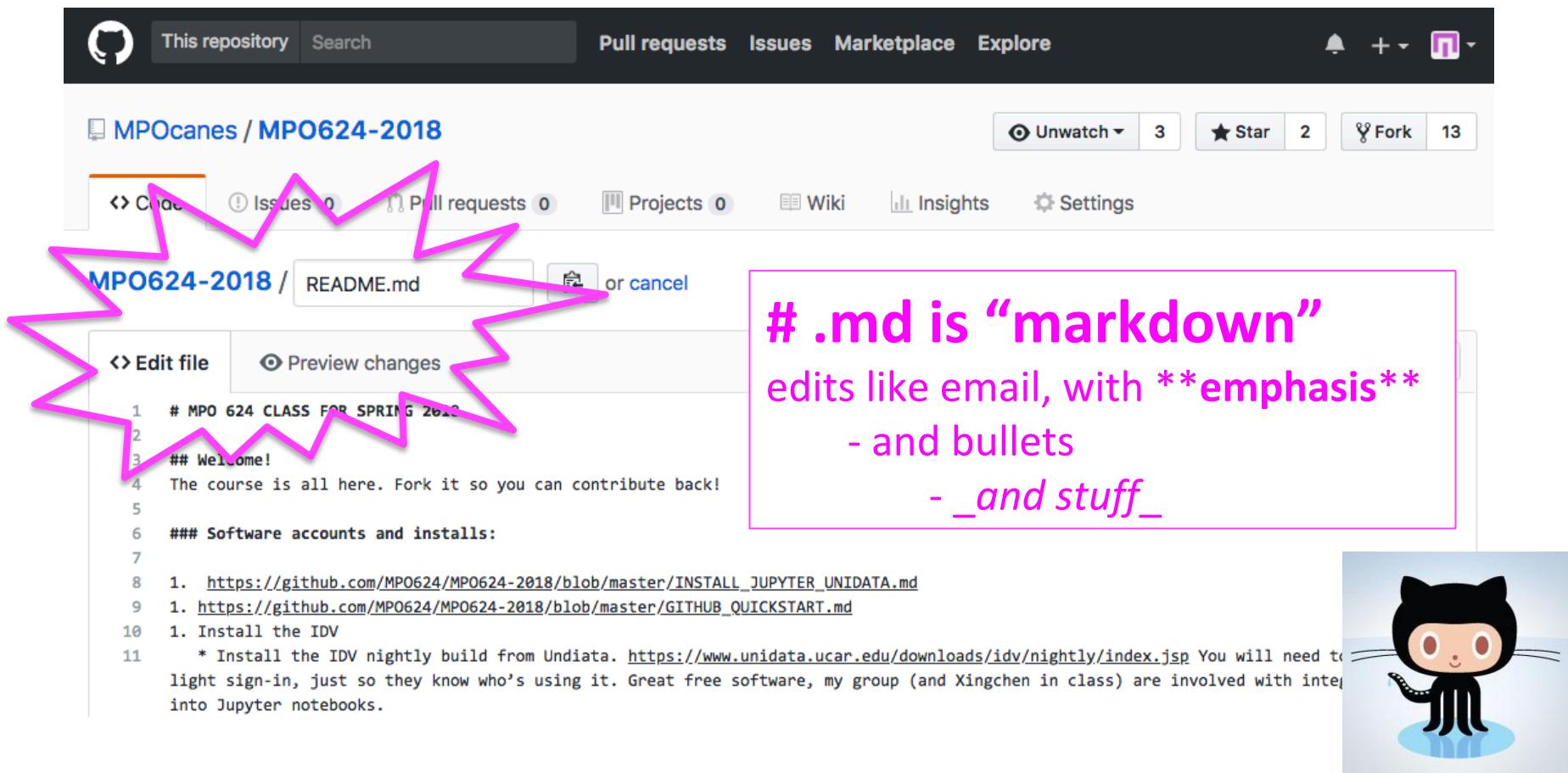
The course is all here. Fork it so you can contribute back!



# conda-jupyter GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world



# .md is “markdown”  
edits like email, with **\*\*emphasis\*\***  
- and bullets  
- and stuff

MPO624-2018 / README.md

```
1 # MPO 624 CLASS FOR SPRING 2018
2 ## Welcome!
3 The course is all here. Fork it so you can contribute back!
4
5 ### Software accounts and installs:
6
7 1. https://github.com/MPO624/MPO624-2018/blob/master/INSTALL\_JUPYTER\_UNIDATA.md
8 1. https://github.com/MPO624/MPO624-2018/blob/master/GITHUB\_QUICKSTART.md
9 1. Install the IDV
10    * Install the IDV nightly build from Undidata. https://www.unidata.ucar.edu/downloads/idv/nightly/index.jsp You will need to light sign-in, just so they know who's using it. Great free software, my group (and Xingchen in class) are involved with integrating it into Jupyter notebooks.
```



conda-juPy-

# GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world

GitPrint – Easily print GitHub markdown

Secure | https://gitprint.com/MPO624-2018/blob/master/README.md

Apps Save to Mendeley NSFC NNC TC tide ClimRean radar DuPage RAP\_wx Unisys INFO AMSj modis CPC Library deG

**Change github to gitprint for a fine looking document! →**

**MPO 624 CLASS FOR SPRING 2018**

---

**Welcome!**

The course is all here. Fork it so you can contribute back!

**Software accounts and installs:**

1. [https://github.com/MPO624/MPO624-2018/blob/master/INSTALL\\_JUPYTER\\_UNIDATA.md](https://github.com/MPO624/MPO624-2018/blob/master/INSTALL_JUPYTER_UNIDATA.md) ([https://github.com/MPO624/MPO624-2018/blob/master/INSTALL\\_JUPYTER\\_UNIDATA.md](https://github.com/MPO624/MPO624-2018/blob/master/INSTALL_JUPYTER_UNIDATA.md))
2. [https://github.com/MPO624/MPO624-2018/blob/master/GITHUB\\_QUICKSTART.md](https://github.com/MPO624/MPO624-2018/blob/master/GITHUB_QUICKSTART.md) ([https://github.com/MPO624/MPO624-2018/blob/master/GITHUB\\_QUICKSTART.md](https://github.com/MPO624/MPO624-2018/blob/master/GITHUB_QUICKSTART.md))

# conda-jupyter-GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world

The screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The page includes a navigation bar with links for 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. A sidebar on the right contains a large pink annotation that reads 'Add files by drag and drop to browser!' with arrows pointing to the 'Create new file', 'Upload files', and 'Find file' buttons. The main content area displays repository statistics: 378 commits, 2 branches, 0 releases, and 13 contributors. It also shows a dropdown for the branch 'master' and a 'New pull request' button. Below the stats, a commit history lists changes by user 'brianmapes' to 'README.md'. The bottom right corner features the GitHub logo (a black octocat) and the text '6 days ago'.

MPOcanes / MPO624-2018

Unwatch

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

MPO 624 CLASS FOR SPRING 2018

atmospheric science idv notebook jupyter-notebook jupyter atmospheric-science oceanograph

378 commits 2 branches 0 releases 13 contributors

Branch: master New pull request Create new file Upload files Find file Clone or download

brianmapes Update README.md

assignments Update README.md

classnotes Create 2018-02-08.md

exams update from MPO624

lessonplans Update 5thTuesday.md

notebooks Add some notebooks

Latest commit 1327307 20 hours ago

6 days ago

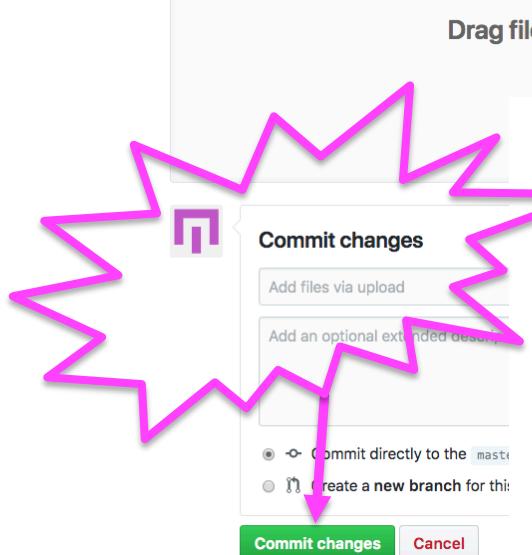
# conda-jupyter-GitHub: tiers of commitment

## 2. Sign up, and create your own *repositories*

Free! If open to whole world

This screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The top navigation bar includes links for 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the bar, the repository name 'MPOcanes / MPO624-2018' is displayed, along with statistics: 3 unwatched, 2 stars, and 13 forks. A large pink arrow points from the text 'Add files by drag and drop to browser!' towards the 'Drag files here to add them to your repository' area.

Add files by  
drag and drop  
to browser!



But understand: a *repository*  
is more than a dump for files!  
every **commit** requires a  
**commit message...**  
it's all logged, that's good



# conda-juPy-GitHub: tiers of commitment

It's why there are 7 layers: acts as a *quality filter*

This screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The page includes a navigation bar with links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the navigation bar, the repository name 'MPOcanes / MPO624-2018' is displayed along with statistics: 3 pull requests, 2 stars, and 13 forks. The main content area shows a commit by 'brianmaps' titled 'Add some notebooks' made 6 days ago. The commit details show 1 contributor and file statistics: 2409 lines (2408 sloc) and 85.2 KB. At the bottom, there are links for 'Raw', 'Blame', and 'History'.

## Jupyter Notebook Users Manual

This page describes the functionality of the [Jupyter](#) electronic document system. Jupyter documents are called "notebooks" and can be seen as many things at once. For example, notebooks allow:

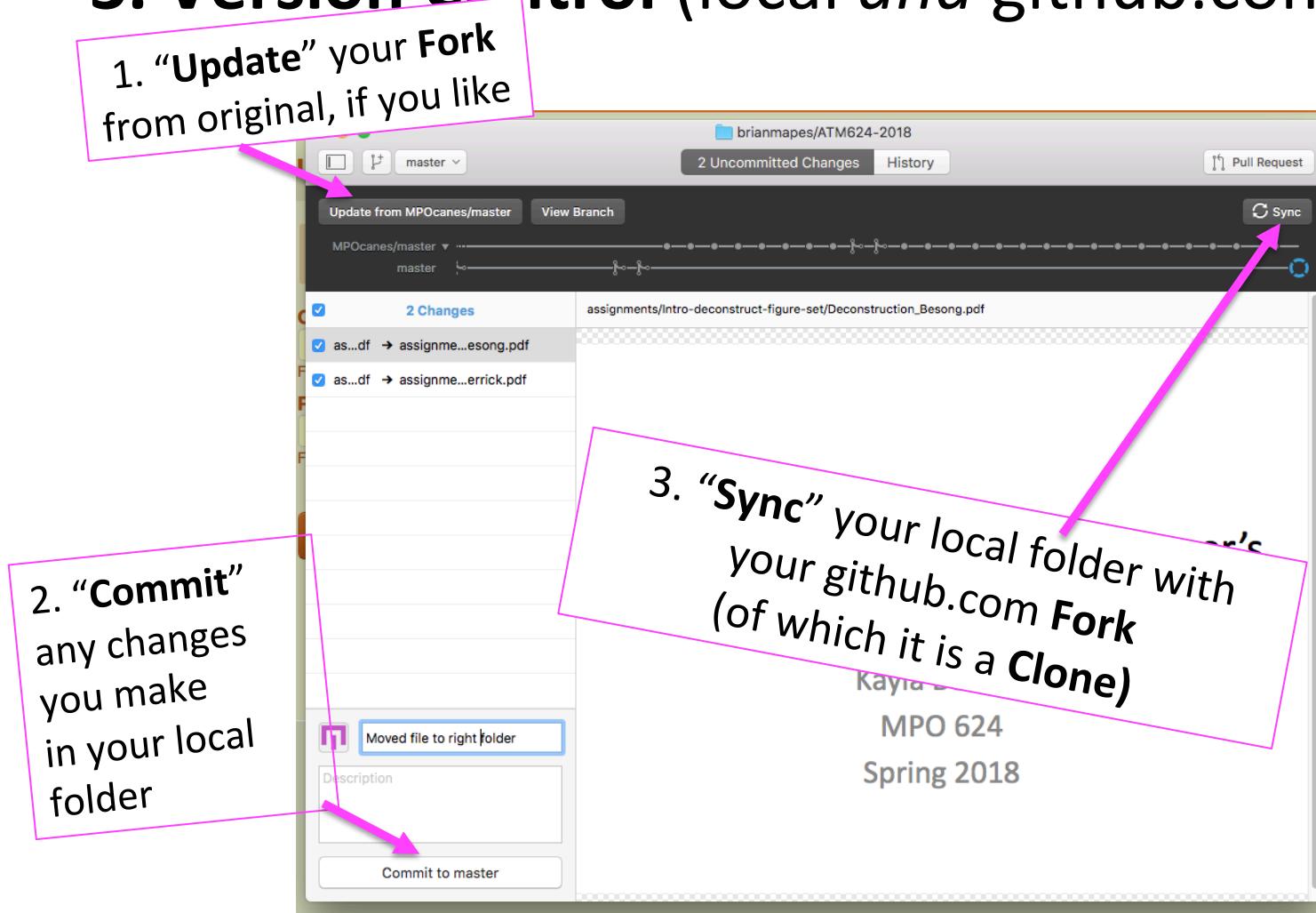
- creation in a **standard web browser**
- direct **sharing**
- using **text with styles** (such as italics and titles) to be explicitly marked using a [wikitext language](#)
- easy creation and display of beautiful **equations**
- creation and execution of interactive embedded **computer programs**
- easy creation and display of **interactive visualizations**

Jupyter notebooks (previously called "IPython notebooks") are thus interesting and useful to different groups of people:

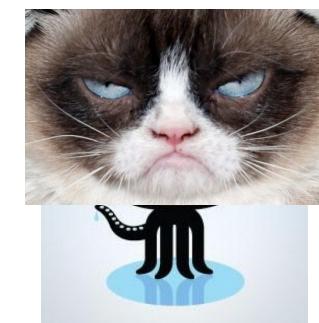


# conda-jupyter GitHub: tiers of commitment

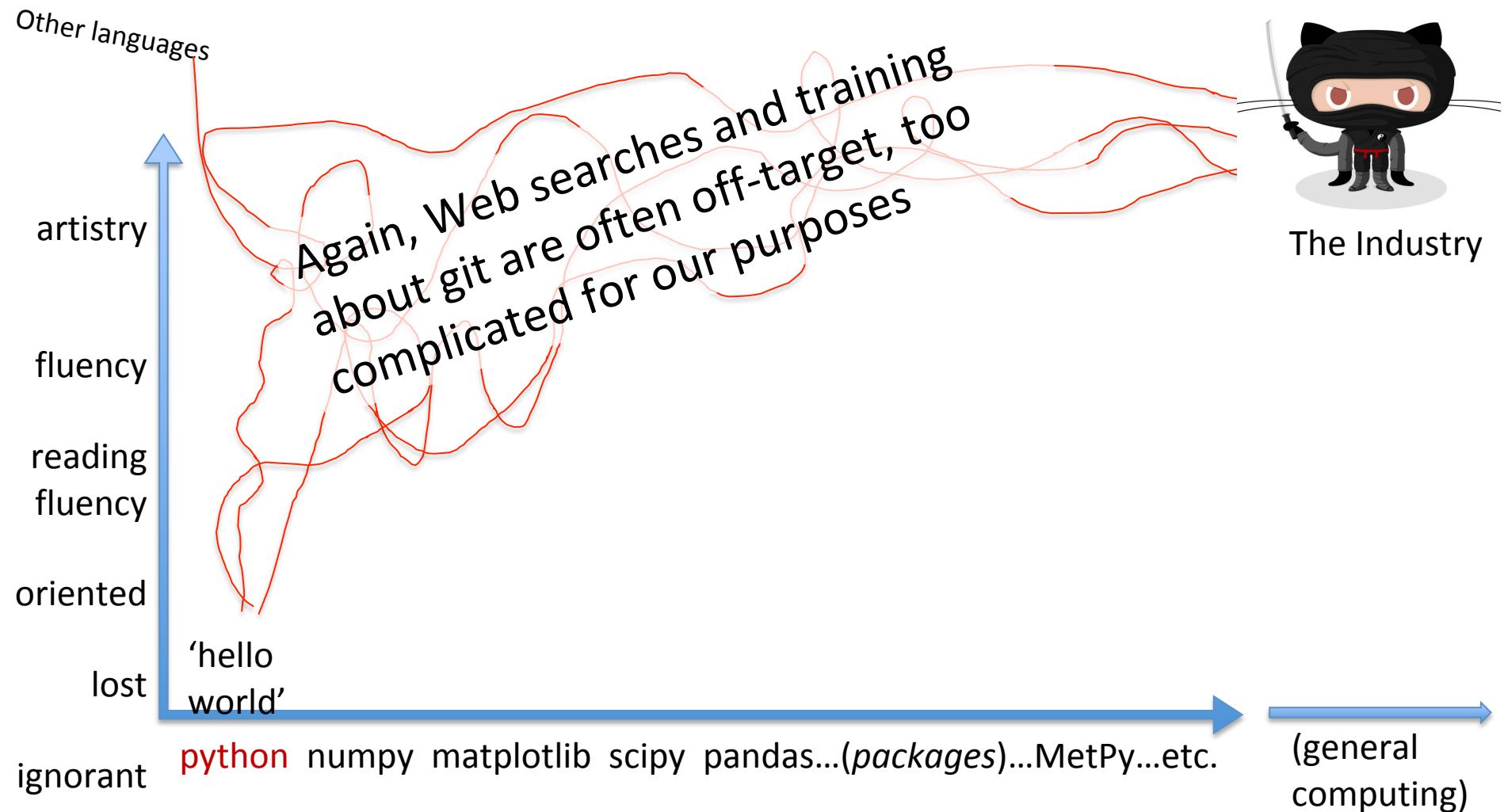
## 3. Version Control (local and github.com backup)



\* Desktop client  
MUUCH nicer  
than memorizing  
git commands!  
*“sync” and  
“update” are not  
git commands!*

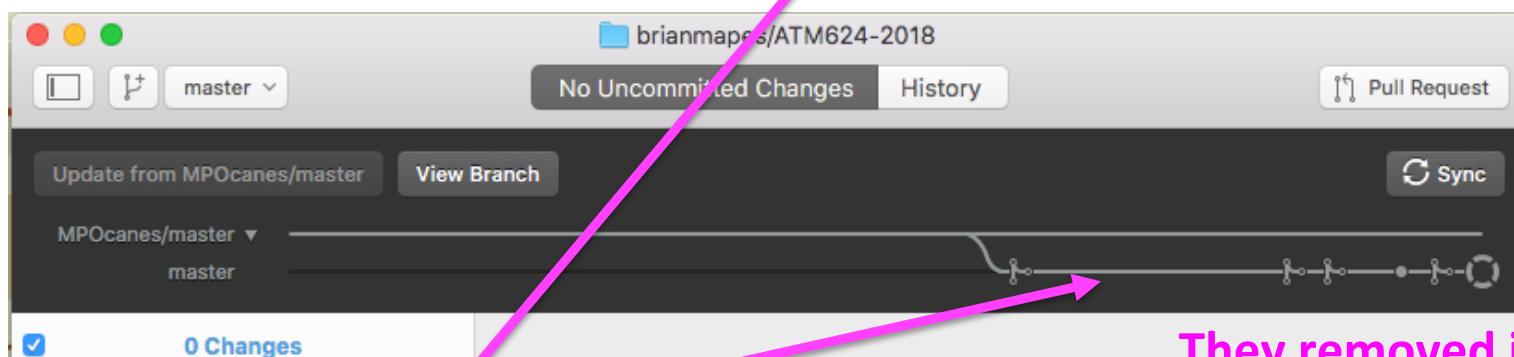
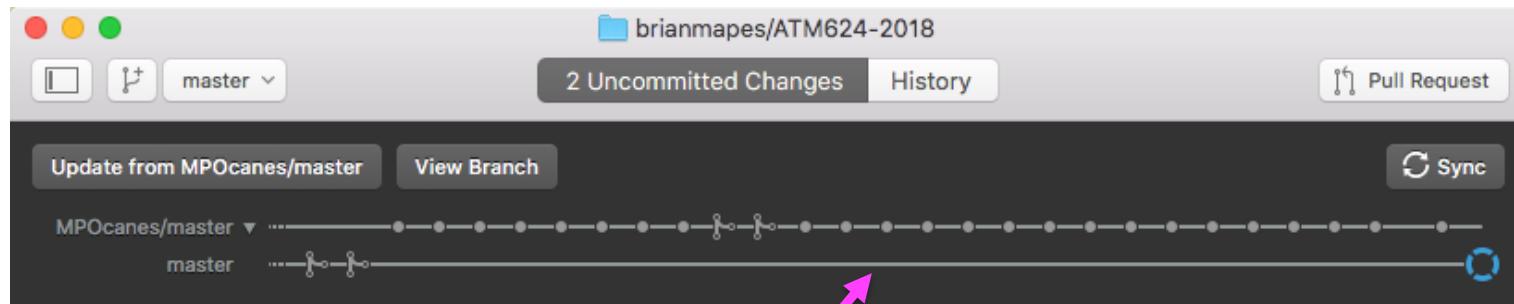


git and github are mainly designed for  
**teams of software engineers**, not for  
occasional-collaborators like us...



conda-jupyter GitHub: tiers of commitment

### 3. Version Control (local and github.com backup)

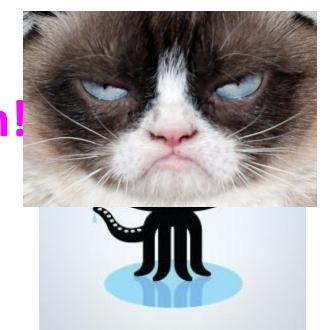


"subway plot" shows great timeline of all the updates, commits, merges!

They removed it from latest version!  
Grrr!

Get old "223" version!

\* Desktop MUUCH nicer than memorizing git commands!  
“sync” and “update” are not git commands!



conda-jupyter GitHub: tiers of commitment

## 4. Collaborate!

The screenshot shows a GitHub repository page for 'MPOcanes / MPO624-2018'. The page includes a navigation bar with links for 'This repository', 'Search', 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the navigation bar, there's a header with the repository name 'MPOcanes / MPO624-2018', a 'Unwatch' button (with 3 notifications), a 'Star' button (with 2 stars), a 'Fork' button (with 13 forks), and a 'FORK ME!' button. The main content area displays repository statistics: 379 commits, 2 branches, 0 releases, and 13 contributors. It also features tabs for 'Code', 'Issues 0', 'Pull requests 0', 'Projects 0', 'Wiki', 'Insights', and 'Settings'. Below these are buttons for 'Edit', 'Manage topics', and a dropdown for 'Branch: master'. A 'New pull request' button is highlighted with a pink arrow. At the bottom, there are buttons for 'Create new file', 'Upload files', 'Find file', and a prominent green 'Clone or download' button. A callout box on the left side contains the text: 'Improved your fork in some generalized way? Send a pull request to the owner!'. To the right of the callout box is a GitHub logo icon. The entire image is annotated with several pink arrows and starburst shapes pointing to specific buttons and text elements.

FORK ME!

more than just

DOWNLOAD

Improved your fork in some generalized way?

Send a **pull request** to the owner!

GitHub logo icon

# Outline

- Appreciating the landscape
  - The IT Stack: hardware, platforms, API ‘levelers’, and user interfaces for scientists
  - Python world, the province of science, and where a web search tends to drop you
- *Why all of conda-juPy-GitHub is worth it*
- **Examples of the goodies out there**
- Where to find **conda-juPy-GitHub onramp**
- Showcase (finish in a notebook)

Unidata / unidata-python-workshop

Code Issues Pull requests Projects

Branch: master unidata-python-workshop / notebooks /

dopplershift Remove use of "lIs" (Fixes #264) ...

..

Animation	Remove solutions from i
Bonus	Move wms sample to fa
CartoPy	Remove solutions from i
Command_Line_Tools	Modify header image ac
GOES_RGB_Demo	MNT: Fix notebooks on
Jupyter_Notebooks	Modify header image ac
Matplotlib	Break up Primer notebo
MetPy_Advanced	Remove unused import
MetPy_Case_Study	Remove use of depreca
Metpy_Introduction	Add stand alone MetPy
Model_Output	Rename model data to r
NumPy	Break up Primer notebo
Primer	Plot magnitude of temp
Pythonic_Data_Analysis	← I showed this one earlier Change buoy data date
Satellite_Data	Fix spelling errors in sat
Siphon	Remove use of "lIs" (Fix
Skew_T	Add solution to bulk she
Surface_Data	Fix surface notebook
Time_Series	Change dates to 9/6-9/
netCDF	Remove use of "lIs" (Fix

# Notebook Troves

- Unidata's educational (training) "boosters":
  - <https://github.com/Unidata/unidata-python-workshop/tree/master/notebooks>



# Unidata's trove

- How-to for technical skills, like:

## Creating Animations

Unidata Python Workshop



### Overview:

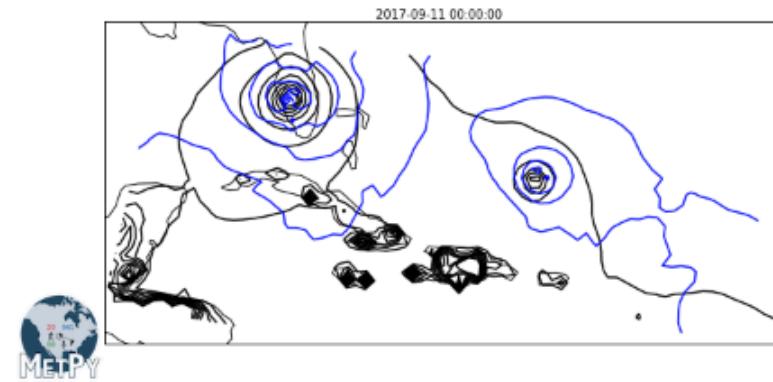
- Teaching: 15 minutes
- Exercises: 15 minutes

### Questions

1. How are animations created with Matplotlib?
2. How can the data in a matplotlib plot be changed?
3. How can animations be displayed in the Jupyter Notebook?

### Objectives

1. [Create a basic animation with matplotlib](#)
2. [Download model output with Siphon](#)
3. [Create an animated contour analysis of the output](#)



Once  Loop  Reflect

# Unidata's flagship package: MetPy

- Software eng-quality tools - for **met!**
  - many calculations, constants, I/O routines, etc.
    - all quantities carry **units**: safe coding!
      - awkward initially, but so is debugging without them...
        - » it grows on you – after learning curve

# MetPy's many powers

- *fetch and plot any sounding from the online Wyoming archive, by station name and date:*

```
from siphon.simplewebservice.wyoming import WyomingUpperAir
from metpy.io import get_upper_air_data

date = datetime(2017, 9, 10, 6)
station = 'MFL'
df = WyomingUpperAir.request_data(date, station)
```

I used it, & offered *my first PR!* (pull request)

Jupyter nbviewer

JUPYTER FAQ </>   

ATM407 / Energy\_Mass\_Plot\_Wyoming\_Hyperlinks.ipynb

- ...to which they responded w/ much education & help...
    - leading to partial acceptance of the most generalizable aspects
      - *all of which they value enough to happily facilitate in a friendly way*

**true collaboration – feels rare and novel, actually**

# Outline

- Appreciating the landscape
  - The IT Stack: hardware, platforms, API ‘levelers’, and user interfaces for scientists
  - Python world, the province of science, and where a web search tends to drop you
- *Why all of conda-jupyter-GitHub is worth it*
- **More examples of the goodies out there**

# The bigger world of Jupyter:



Kalman-and-Bayesian-Filters-in-Python / [table\\_of\\_contents.ipynb](#)



## Kalman and Bayesian Filters in Python

*Table of Contents*

[Preface](#)

Motivation behind writing the book. How to download and read the book. Requirements for IPython Notebook and Python. github links.

[Chapter 1: The g-h Filter](#)

Intuitive introduction to the g-h filter, also known as the  $\alpha$ - $\beta$  Filter, which is a family of filters that includes the Kalman filter. Once you understand this chapter you will understand the concepts behind the Kalman filter.

[Chapter 2: The Discrete Bayes Filter](#)

**Executable Books!**



## Earth Science and Geo-Spatial data

- [EarthPy](#), a collection of IPython notebooks with a focus on Earth Sciences, from [whale tracks](#) to the [flow of the Amazon](#).
- Python for Geosciences, a tutorial series aimed at the Earth Sciences community, by [Nikolay Koldunov](#).
- Find graffiti close to NY subway entrances, one of a rich [collection of notebooks](#) on large-scale data analysis, by [Roy Hyunjin Han](#).
- Logistic models of well switching in Bangladesh, part of the "Will it Python" blog series ([repo](#)) on Machine Learning and data analysis in Python. By Carl Vogel.
- Estimated likelihood of observing a large earthquake on a continental low-angle normal fault and implications for low-angle normal fault activity, an executable version of a paper by Richard Styron and Eric Hetland published in *Geophysical Research Letters*, on earthquake probabilities.
- [python4oceanographers](#), a blog demonstrating analyses in physical oceanography from resource-demanding numerical computations with functions in compiled languages to specialized tidal analysis to visualization of various geo data using fancy things like interactive maps.

**Executable Papers!**

## Reproducible academic publications

---

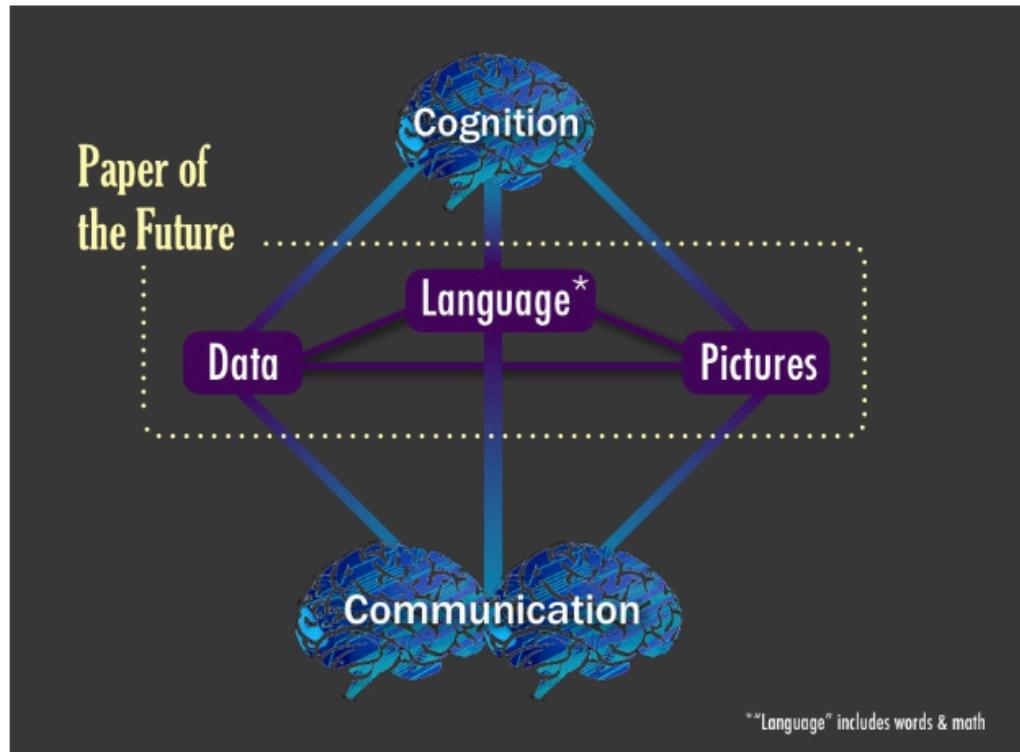
This section contains academic papers that have been published in the peer-reviewed literature or pre-print sites such as the [ArXiv](#) that include one or more notebooks that enable (even if only partially) readers to reproduce the results of the publication. If you include a publication here, please link to the journal article as well as providing the nbviewer notebook link (and any other relevant resources associated with the paper).

1. [Discovery of Gravitational Waves by the LIGO collaboration](#). That page, from the LIGO Open Science Center, contains multiple notebooks for various datasets corresponding to different events; [this binder](#) lets you run the code right away. More details on the [GW150914](#) event as well as the original [main Physical Review Letters paper](#), "Observation of Gravitational Waves from a Binary Black Hole Merger".
2. [Characterizing Strain Variation in Engineered E. coli Using a Multi-Omics-Based Workflow](#), by Brunk et al.
3. [Predicting Coronal Mass Ejections Using Machine Learning Methods](#) by Monica Bobra and Stathis Ilonidis (Astrophysical Journal, 2016). An [IPython notebook](#) which reproduces all the results, has been permanently deposited in the [Stanford Repository](#).
4. [The Paper of the Future](#) by A. H. G. de Jong et al. (arXiv:1701.02413, 2017). This article explains and shows with detailed examples how the future of scientific "papers" can morph into long-lasting non-records of scientific discourse, enriched with deep data and code linkages, interactive figures, audio, video, and commenting. It includes an interactive d3.js visualization and has an astronomical data figure with an IPYthon Notebook "behind" it.

A Movement...

scenarios, and to explain findings more clearly. **This short article explains—and shows with demonstrations—how scholarly "papers" can morph into long-lasting rich records of scientific discourse**, enriched with deep data and code linkages, interactive figures, audio, video, and commenting.

A Movement...



literate programming

**Fig. 1**

*The Paper of the Future should include seamless linkages amongst **data**, **pictures**, and **language**, where "language" includes both words and math. When an individual attempts to understand each of these kinds of information, different cognitive functions are utilized: communication is inefficient if the channel is restricted primarily to language, without easy interconnection to data and pictures.*

## 2 Collaborative Authoring

# Our whole curriculum?

## Students do the learning curve *once*

The screenshot shows a GitHub organization page for "MPOcanes". The top navigation bar includes links for "This organization", "Search", "Pull requests", "Issues", "Marketplace", "Explore", and user notifications. Below the header, the organization's logo (a purple stylized "M") and name "MPOcanes" are displayed. A navigation bar below the logo offers links to "Repositories 1", "People 1", "Teams 0", "Projects 0", and "Settings". Search and filter options are available, along with a "New" button to create a new repository. The main content area features a repository card for "MPO624-2018", which is described as "MPO 624 CLASS FOR SPRING 2018". The repository has tags for "science", "jupyter", "notebook", "jupyter-notebook", and "oceanography". It was created by "Jupyter Notebook" 2 hours ago, has 2 stars, 13 forks, and was updated 2 hours ago. To the right of the repository card are two sidebar boxes: "Top languages" (showing Jupyter Notebook) and "People" (listing "brianmapes" with the name "Brian Mapes" and a "Invite someone" button).

MPOcanes

Repositories 1 People 1 Teams 0 Projects 0 Settings

Search repositories... Type: All Language: All Customize pinned repositories New

**MPO624-2018**  
MPO 624 CLASS FOR SPRING 2018

science jupyter notebook jupyter-notebook oceanography

● Jupyter Notebook ★ 2 13 Updated 2 hours ago

Top languages

● Jupyter Notebook

People

brianmapes Brian Mapes

Invite someone

# Our whole curriculum?

3-4 classes so far

The screenshot shows the GitHub interface for the organization "ATMocanes". The top navigation bar includes links for "This organization", "Search", "Pull requests", "Issues", "Marketplace", "Explore", and a notifications icon. Below the header, the organization's logo (a purple stylized 'T') and name "ATMocanes" are displayed. A navigation bar below the logo offers links to "Repositories 3", "People 1", "Teams 0", "Projects 0", and "Settings". Search and filter options ("Search repositories...", "Type: All", "Language: All") are also present. On the left, three repository cards are listed: "ATM406" (MIT license, updated 16 seconds ago), "ATM303" (Jupyter Notebook, MIT license, updated 3 days ago), and "ATM407" (Jupyter Notebook, 1 update, updated on Jan 8). To the right, two sidebar boxes are visible: "Top languages" (Jupyter Notebook) and "People" (listing "brianmapes" and "Brian Mapes" with an "Invite someone" button).

This organization Search Pull requests Issues Marketplace Explore

ATMocanes

Repositories 3 People 1 Teams 0 Projects 0 Settings

Search repositories... Type: All Language: All Customize pinned repositories New

ATM406

MIT Updated 16 seconds ago

ATM303

Jupyter Notebook MIT Updated 3 days ago

ATM407

Jupyter Notebook 1 Updated on Jan 8

Top languages

Jupyter Notebook

People

brianmapes Brian Mapes

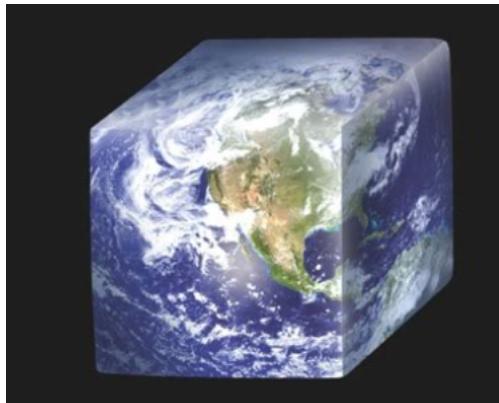
Invite someone

My own jupyter-based project...

# DRILSDOWN

## Seeing the Interesting Parts of Big Data

Brian Mapes, Suvarchal Cheedela,  
Mo Iskandarani, Xingchen Yang (Miami)  
Yuan Ho, Cecelia Hedrick, others (Unidata)  
Jeff McWhirter (Geode Systems)



An NSF Earthcube  
Building Blocks project



# DRILSDOWN

IDV  
“bundles”

- Drawing
- Rich
- Integrated
- Lat-lon-time
- Subsamples, from
  - Datasets
  - Online, into
- Working
- Notebooks

“Teleport”

Aggregations in  
Repositories

jupyter

# IDV in jupyter: Case Studies are our currency good for synoptic lab – and more...

Sandy\_case\_study

localhost:8888/notebooks/DRILSDOWN\_DEMO/Sandy\_case\_study.ipynb#Anchor3

jupyter Sandy\_case\_study Last Checkpoint: 8 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python [default]

In [5]: `*make_image -caption 'upper_and_lower_vortices_nearing_interaction'`

Vorticity in 3D, the 500 mb height footprint, and clouds & precip:

- Tell the story

Time 2012-10-28 00:00:00 UTC

'upper\_and\_lower\_vortices\_nearing\_interaction'

Caption: An upper trough and hurricane Sandy in Absolute Vorticity and Z500 fields

# Outline

- Appreciating the landscape
  - The IT Stack: hardware, platforms, API ‘levelers’, and user interfaces for scientists
  - Python world, the province of science, and where a web search tends to drop you
- *Why all of conda-juPy-GitHub is worth it*
- Where to find **conda-juPy-GitHub onramp**
- Showcase (finish in a notebook)

<https://github.com/MPOcanes/MPO624-2018>

This repository Search Pull requests Issues Marketplace Explore

MPOcanes / MPO624-2018 Onramp Unwatch 3 Star 2 Fork 13

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

MPO 624 CLASS FOR SPRING 2018 Edit

atmospheric science idv notebook jupyter-notebook jupyter atmospheric-science oceanography unidata Manage topics

379 commits 2 branches 0 releases 13 contributors

Branch: master New pull request Create new file Upload files Find file Clone or download

brianmaps Update README.md Latest commit 07990e9 2 hours ago

assignments Update README.md : ago

classnotes Create 2018-02-08.md : ago

exams : ago

... assignment 1 : ago

Update CALENDAR.md : ago

Create CONDA\_ENVIRONMENTS.md : ago

Merge remote-tracking branch 'MPO624/master' 15 days ago

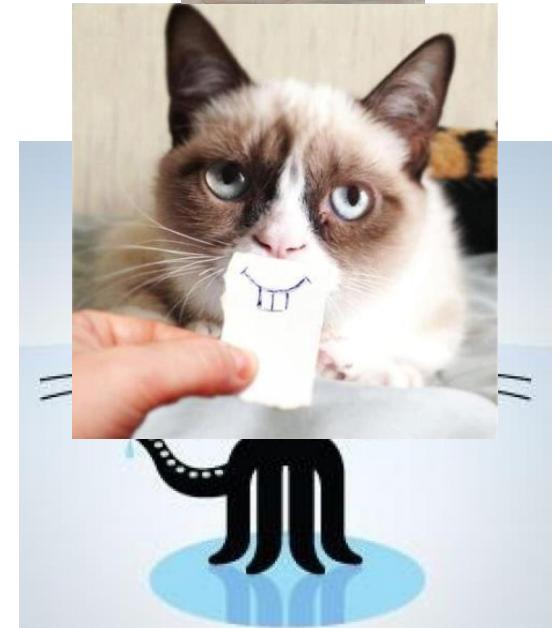
Update INSTALL\_JUPYTER\_UNIDATA.md 3 days ago

CONDA\_ENVIRONMENTS.md  
GITHUB\_QUICKSTART.md  
INSTALL\_JUPYTER\_UNIDATA.md

I've honed it down pretty clean and minimal for MPO624...

# Conclusion: conda-juPy-GitHub is tolerable to learn, to unlock...

- ...so much free goodness!
  - unification of computing, results, & communication
  - work from troves of examples
    - #NeverAgainStartFromScratch
  - true collaboration
    - with discoverable strangers
- But, know your mavens
  - Search is an ally – not quite friend
    - works for industry, mostly



# And – it's a slide show too!

- (demo RISE package, reveal.js)

A screenshot of a Jupyter Notebook interface. The top bar shows the title "Lecture\_Siqueira\_distributions-py3", the status "Last Checkpoint: a day ago (autosaved)", and the kernel information "Not Trusted | Python 3 C". A pink arrow points from the text above to the "Enter/Exit Live Reveal Slideshow" button in the toolbar. The main content area displays a slide with the title "MPO 624: Histograms and PDFs", contributed by Leo Siqueira, and a "Steps" section listing "1. [Histograms](#)" and "2. [Kernels](#)". The slide footer features the University of Miami logo.

jupyter Lecture\_Siqueira\_distributions-py3 Last Checkpoint: a day ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3 C

Enter/Exit Live Reveal Slideshow

MPO 624: Histograms and PDFs

Contributed by Leo Siqueira

---

Steps

1. [Histograms](#)
2. [Kernels](#)

Slide Type Slide

MIAMI