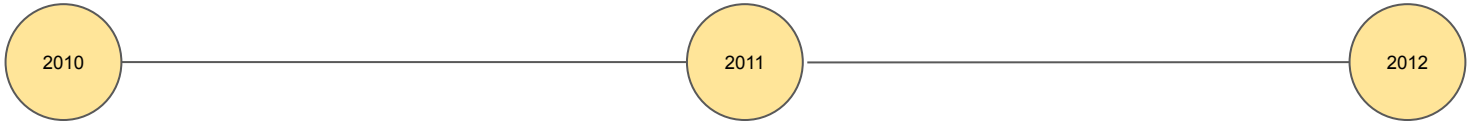
OSS Contributions: An external perspective, from a long time ago…

Phil Elson
Beams dept & OSPO co-representative
21/02/2025

Open source is a **long-term** approach

For software that needs to evolve with the times, I believe it offers the best return on investment over the long-term



Working in “Data, Analysis and Visualisation” team at UK’s national weather & climate lab: Met Office

Scientific Python is beginning to look like an interesting option. Numpy at v1.2, matplotlib (visualisation tool) at v0.99.

Formal Objectives:

- Match the existing map drawing capabilities (of IDL/PV-WAVE) in Python
- Become the org’s specialist in matplotlib

Exploring the capabilities of existing mapping library

First contact with the matplotlib community mailing list

- Providing clarity on what I was trying to achieve, and confirming limitations of existing functionality

First pull request (PR) showing an end-to-end working prototype

First PR merged (>6 months after opening)

Additional PRs adding small functionality required

Start contributing to broader design discussions, and answering issues on the tracker

Additional PRs which refractor and simplify some of the more complex parts of the codebase

Given “commit bit”, aka. merge rights

A long time ago...

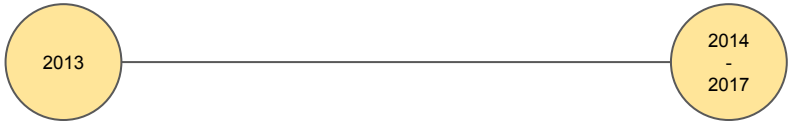
I was promoted relatively quickly after first major matplotlib contributions

Not because *I was special*, but because the *project leader was special*:

- The project's view:

  You get more by promoting *good people early* than you lose by promoting the *wrong people too soon**

* It may have been luck that the project never really had the wrong people.
  A project needs to make its expectations of conduct clear.
  Dealing with bad behaviour is hard, but it is even harder if the definition of "bad" is undefined



First in-person meeting with collaborators I spent a lot of time working with

Instant common interests -> long-lasting friendships

Introduced to fellow project maintainers (e.g. numpy, scipy, IPython/Jupyter, …)

Meeting was a platform to launch a *moderately successful* mapping library built on top of matplotlib

Meeting peers, friends and fellow maintainers while travelling: in places ranging from from Australia to Hawaii

Continuous ~30% effort on quality of life improvements for matplotlib, code reviews and issue triage (both personal and work time)

Joined the steering committee for the project:

- Fostering community
- Contributing to high-level design
- Coordinating conference efforts
- …

Not always working on stuff that my org needed!
I was contributing to the wellbeing of the project (protecting the investment)

This was part of my objectives

Reflection: Was this a good investment?

- Out of the box powerful functionality (SciPy stack!), which had no upfront development effort, nor ongoing license cost
- Specialised functionality (for a Met agency) developed as OSS
  - **Success was aided by the network effect that contributing to matplotlib had**
  - Used outside of the originating organisation, and is now being maintained and enhanced by many organisations
    - Including adaptations (by others) to work out-of-core and distributed across machines - hugely beneficial!
- Great recruitment advertising, and provides motivating work for retention
- The few that moved on typically got good jobs (in OSS) due to in no small part to their visibility and OSS connections

https://matplotlib.org/stable/gallery/showcase/xkcd.html

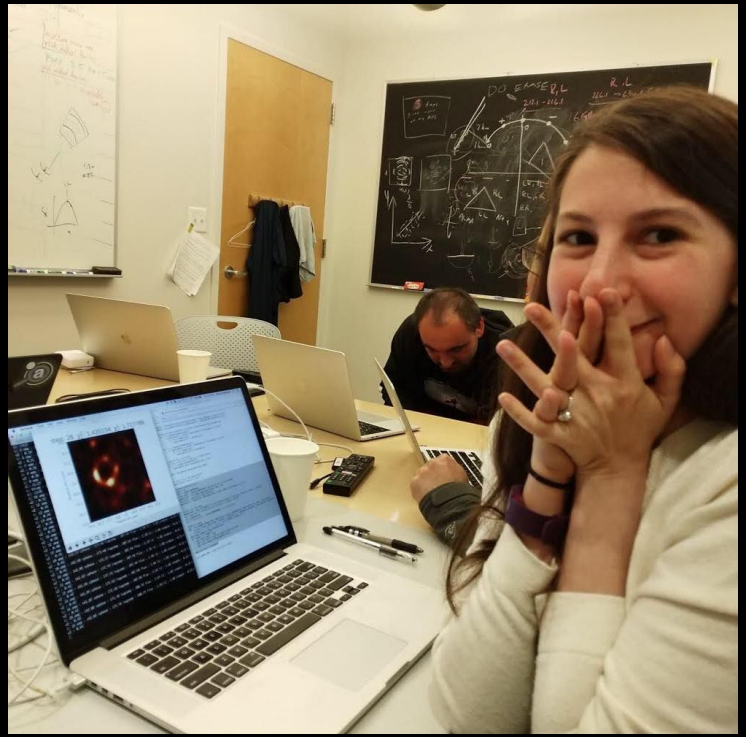
Reflection: Was this a good investment?

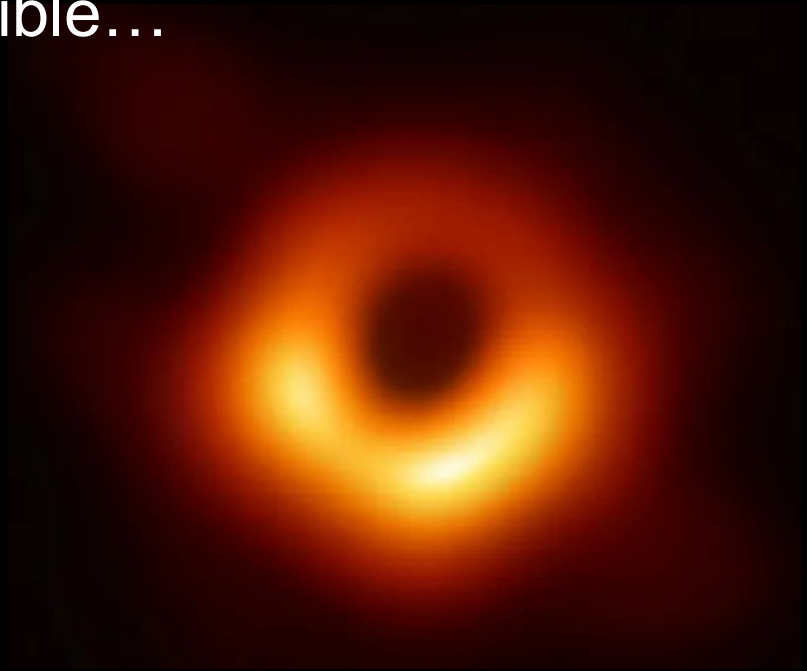
- Not at 100% efficiency though: I didn't *exclusively* work on my organisation's direct needs…

  For example, working on matplotlib's XKCD functionality

"Stove Ownership" from xkcd by Randall Munroe

Impact? Both broad and visible…

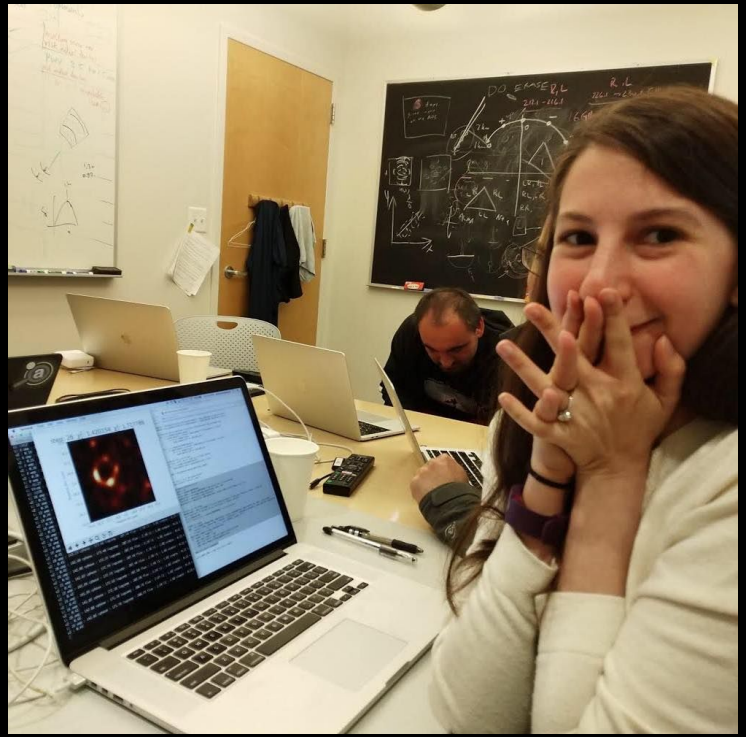




https://edition.cnn.com/2019/04/10/us/katie-bouman-mit-black-hole-algorithm-sci-trnd/index.html

Photo: EHT Collaboration

Impact?

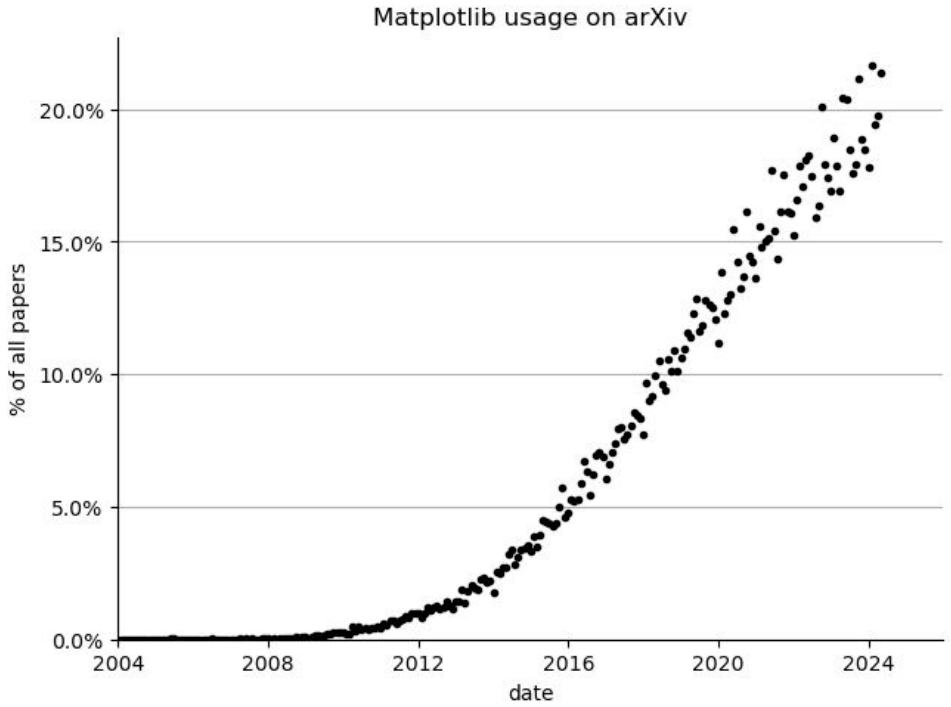


Even from this grainy image, I can see the parts of matplotlib that I contributed to…

https://edition.cnn.com/2019/04/10/us/katie-bouman-mit-black-hole-algorithm-sci-trnd/index.html

Impact?

20% of **all papers** on arXiv used matplotlib



Source: https://docs.coiled.io/user_guide/arxiv-matplotlib.html

Cookbook for continued contributions to a specific project

1. **Understand & research:** find out more about the functionality, as well as how the project is run. Be a user.
2. **Reach out:** confirm your understandings, and then ask whether your specific contributions are welcome
3. **Start small:** it takes time to meet the expectations of an established project
4. **Follow-up:** On reviews, and also *after* your contribution is merged
5. **Join the community:** Help out with the day-to-day activities of the project (e.g. code review, answering questions), and attend events - the network you build is important!

Open source contributions today

- Over time my contributions to matplotlib have wound down. I still love to use it, and make occasionally make a contribution (however there is a strange pressure when contributing to something that you were once an expert at, but which for now you are effectively a novice again)
- I’ve had a work (CERN) objective for a few years around ensuring the longevity of a critical (non-CERN) OSS project which is in a poorly maintained state (work ongoing)
- Generally focussed on small enhancements around core Python (incl. CPython) and packaging communities - less systematically targeting a single project. I have support from my management to work on these topics.

Contributing to OSS opens doors…

- My story of OSS contribution, and ultimately maintainership, leads directly to the creation of a successful OSS project (a story for another time maybe)
- IMO the ROI for an organisation is unbeatable (especially vs creating something new)
    - It is always best to do the work with the support of the organisation (e.g. as a written objective). If not, beware of burnout!
- The visibility can be great for your career
- I value the lasting friendships that I have with fellow contributors from around the world
- The impact of small improvements to widely used projects is huge! It is hard to imagine such humble improvements having such impact across *all* of science
- As an OSPO representative for BE, I'm focussed on clearly documenting CERN's open-source policies, and where possible simplifying procedures around open-source

I personally believe that CERN's societal impact can be enhanced by open-source contributions, and that the effect is maximised when we contribute to projects with established communities