

# Technical Oversharing: Going beyond `Methods` section

Sharing technical knowledge outside research papers is  
good for the environment and good for you

Andrey Andreev, PhD

Bio & Biomedical Engineering dept

[aandreev@caltech.edu](mailto:aandreev@caltech.edu)

<https://aandreev.net>

Updated: Aug 31, 2022

# What is `Technical note`?

From solutions to extra-specialized problems, to community-generated opinions

## How to build & manage 250TB storage server?

### Practical Guide to Storage of Large Amounts of Microscopy Data

Andrey Andreev<sup>1\*</sup> and Daniel E.S. Koo<sup>2</sup>

<sup>1</sup>California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125

<sup>2</sup>University of Southern California, USC Michelson Center for Convergent Bioscience, 1002 West Childs Way, Los Angeles, CA 90089

\*aandreev@caltech.edu

**Abstract:** Biological imaging tools continue to increase in speed, scale, and resolution, often resulting in the collection of gigabytes or even terabytes of data in a single experiment. In comparison, the ability of research laboratories to store and manage this data is lagging greatly. This leads to limits on the collection of valuable data and slows data analysis and research progress. Here we review common ways researchers store data and outline the drawbacks and benefits of each method. We also offer a blueprint and budget estimation for

modern microscopy tools. While our experiments are based on light sheet microscopy, the system described here can be used for any type of image data collection and storage.

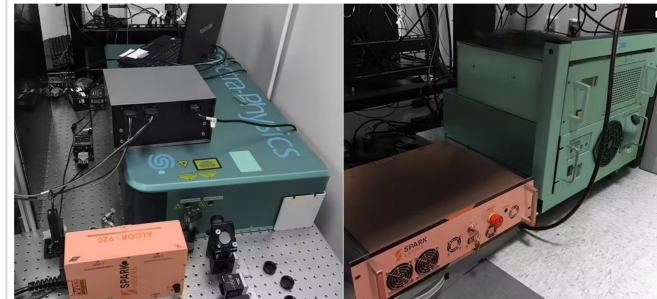
Similar to other microscopy-oriented labs, we collect large amounts of data while simultaneously developing new experiments and data processing pipelines. Much of our work is in constant flux, and it is nearly impossible to reliably lock in a

## Which laser to buy for 2P microscopy?

### Two-photon microscopy and the \$100,000 question: tunable or fixed-wavelength femtosecond laser?

Many biologists are unfamiliar with ultrafast laser technology, but still need to decide between tunable or fixed-wavelength lasers for their two-photon microscopy applications.

Nov. 17, 2020



## What are the benefits and caveats of using CAD to design hardware?

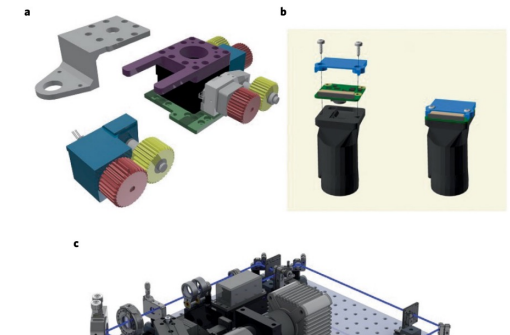
### CAD we share? Publishing reproducible microscope hardware

Here we discuss barriers to reproducibility in regard to microscopes and related hardware, along with best practices for sharing novel designs created using computer-aided design (CAD). We hope to start a fruitful community discussion on how instrument development, especially in microscopy, can become more open and reproducible, ultimately leading to better, more trustworthy science.

Benedict Diederich, Caroline Müllenbroich, Nikita Vladimirov, Richard Bowman, Julian Stirling, Emmanuel G. Reynaud and Andrey Andreev

Microscopy has often been at the heart of new biological discoveries, and as cutting-edge experiments become more complex, so do the new microscopes required to image them. This has led to rapid growth in interdisciplinary collaborations to develop novel instruments. Increasingly, this has been done using reusable building blocks.

Many scientists successfully use CAD software to design custom parts, assemble setups from their components and render graphics. However, these detailed and reusable designs are rarely openly archived. In recent years, funders and publishers have formulated extensive guidelines and policies on the sharing of data and software in the context of open science<sup>1,2</sup>. In contrast, hardware development and its publications often lack suitable standards to ensure the completeness and quality of designs that are shared, even though such standards



And few others, see:  
<https://aandreev.net>

Where? Microscopy Today, LaserFocus World, eLife, Nature Methods, Focal Plane...

# Sharing technical details beyond academic papers: good for the environment, good for you

A lot of technical information is not being written down (published)

- PhD students (like me) want to show-off more but don't have enough time/material to write more whole papers
- Personal blogs are not persistent
- "Putting things on GitHub" is not polished

**Publishing protocols, technical reports and notes can help**

# Sharing technical details beyond academic papers: good for the environment, good for you

Slides available at <https://aandreev.net/tech/>

## Good for the environment:

- Every PhD student once spent months figuring something out; mostly that knowledge is lost
- Many groups face the same problems, reinventing the wheel
- Some technical problems seem easy, thus forcing people on a chase for a solution; working solution might be much more nuanced

## Good for you:

- Scientists are writers; writing is trained by writing more
- Exposure in industry & academia; tangible demonstration of knowledge, skill, and finishing things
- Potential for collaborating with people you wouldn't normally (competitors)

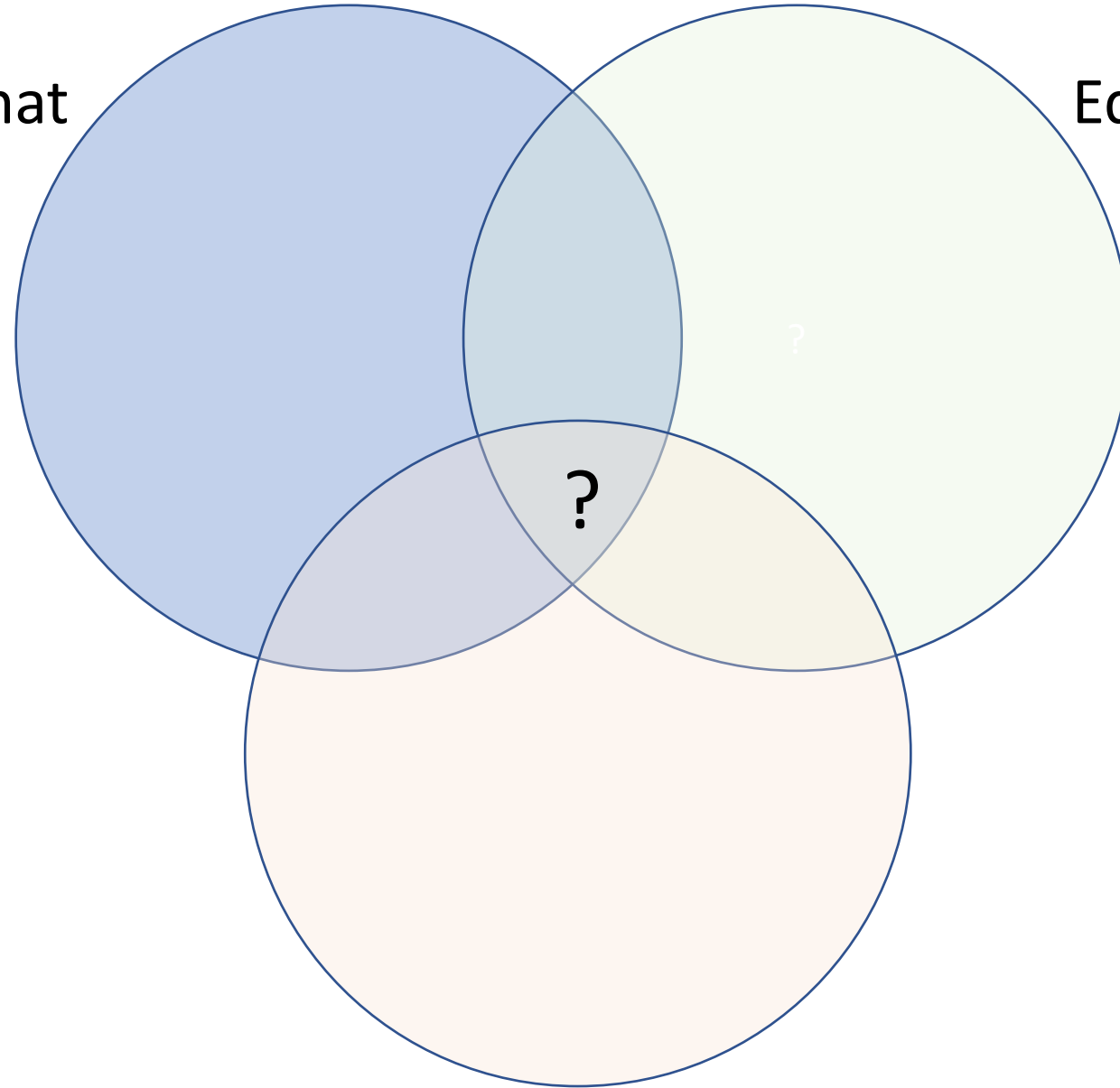
# How to implement more technical writing?

- Identify problem that you know how to solve
  - a lot of time is wasted on solving boring stuff; don't think "its too uninteresting"
- Imagine your audience: what do they care about?
- Clearly identify problem you are solving and audience you are having in your mind
- Don't try to show what you know: try to solve your reader's problem
- My personal inspiration: Larry McEnerney (U Chicago) "functional writing"
  - <https://www.youtube.com/watch?v=vtIzMaLkCaM>
  - [https://www.dropbox.com/s/pk68nxrvvg8uelo8/UnivChic\\_WritingProg-1grt232.pdf](https://www.dropbox.com/s/pk68nxrvvg8uelo8/UnivChic_WritingProg-1grt232.pdf)

But how do I publish it?..

Flexible format

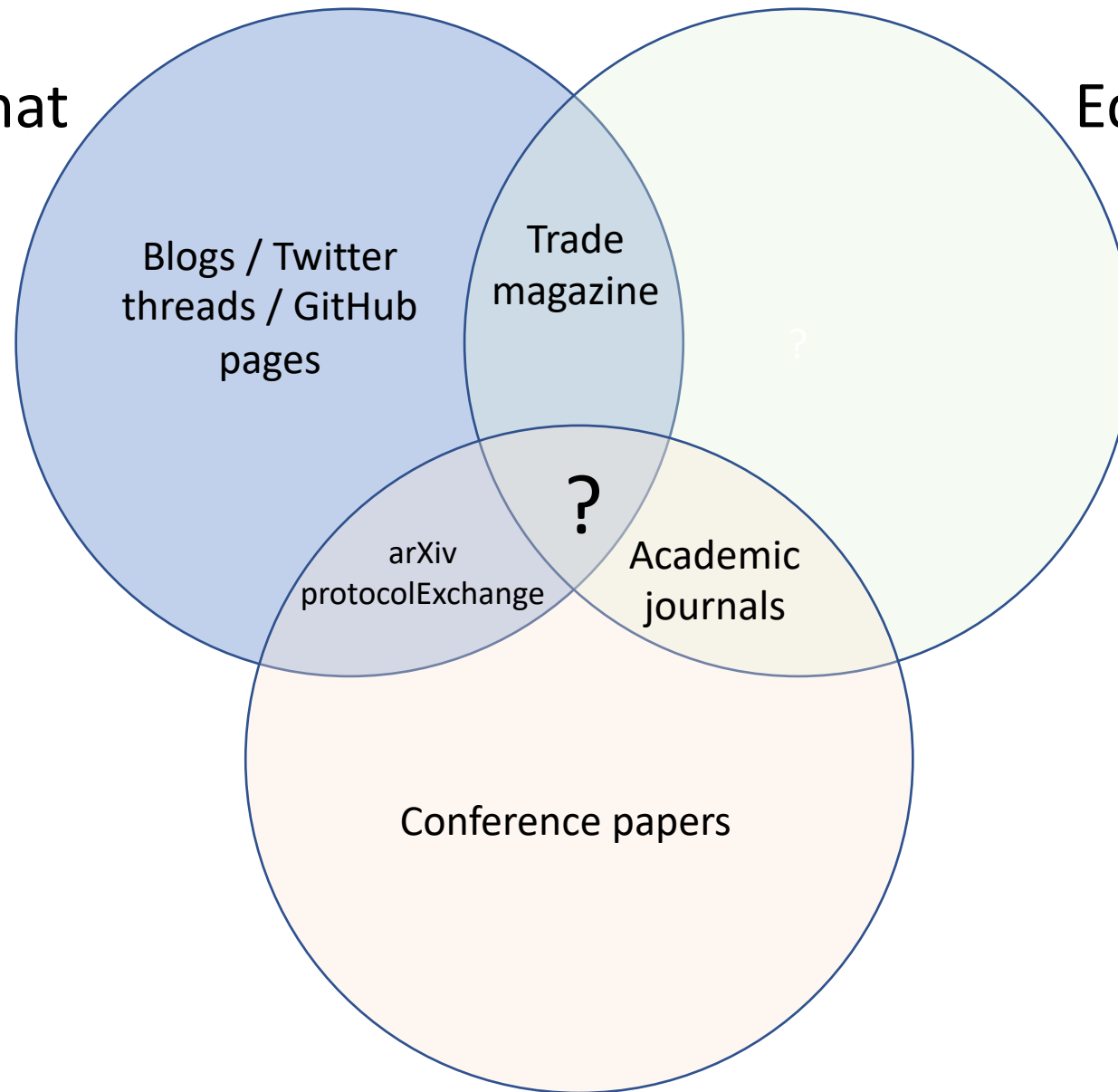
Editorial service



Persistent & Cite-able

Flexible format

Editorial service



Persistent & Cite-able

## Flexible format

For example, blogs from people:

- [labrigger.com](http://labrigger.com)
- [www.michaelchimenti.com](http://www.michaelchimenti.com)
- [www.rp-photonics.com/encyclopedia.html](http://www.rp-photonics.com/encyclopedia.html)
- [www.extremesciencing.org](http://www.extremesciencing.org)
- [www.austinblanco.com/blog](http://www.austinblanco.com/blog)
- [gcamp6f.com/](http://gcamp6f.com/)
- [github.com/AndrewGYork](https://github.com/AndrewGYork)

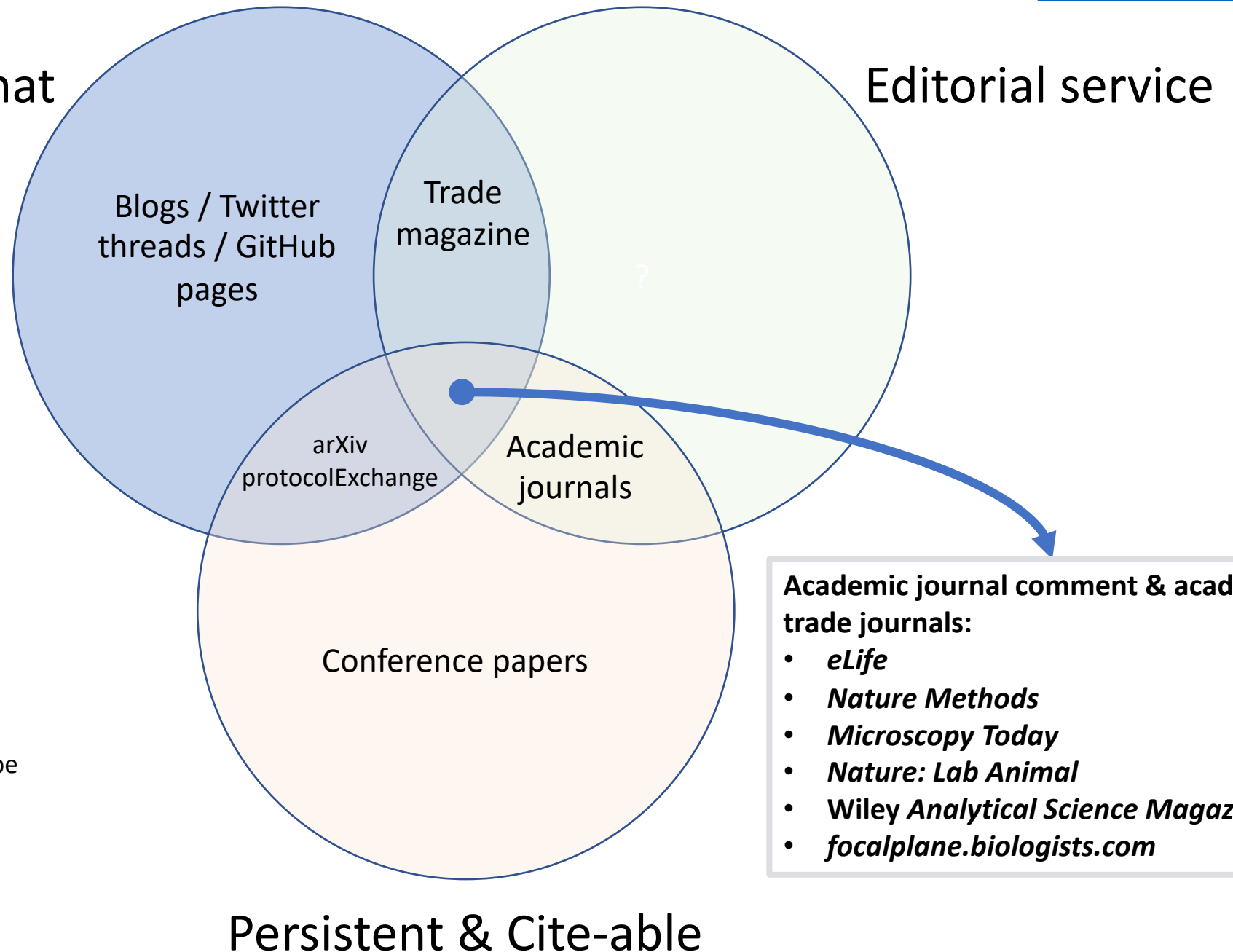
...and companies:

- [www.sptlabtech.com/blog](http://www.sptlabtech.com/blog)
- [blog.addgene.org](http://blog.addgene.org)
- [thorlabs.com](http://thorlabs.com)

**However**, academic system does not incentivize such sharing

Accepted scientific artifacts are limited to peer-reviewed scientific papers (and maybe conference papers e.g. SIGGRAPH)

There are no “reputable” journals for technical notes – hard to use for grants/fellowships



## Persistent & Cite-able



# It might be hard to come up with a topic/subject

## Few potential topics (*that might sound like buzzfeed...*):

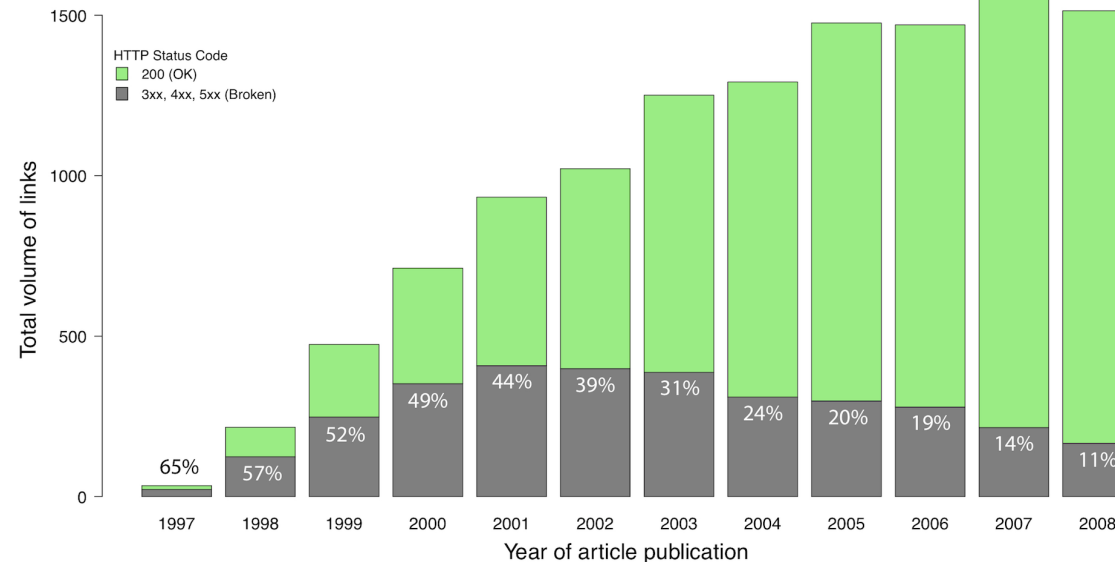
- 5 practical tips for optical alignment
  - Based on existing 25+ twitter threads that ,might disappear in 1 year
- Moving to a new animal facility? Don't forget to buy these 10 things
- Necessary tools for starting optical lab space
- "I got used to {X, Y, Z} but due to supply chain had to sub with {A, B, C}"
- [Software] takes 8 hours to install: Here is step-by-step guide & a virtual machine / Dockers container hosted at CaltechDATA
- Want to analyze/visualize your data using [method/software X]?
  - Here is a walk-through, Jupyter notebooks, and sample dataset to show you the ropes
- Why we will never buy [piece of equipment] / use [method] again
- 10 lab practices that are actually pointless
  - E.g. hold PCR machine at 4C; use RNAzap everywhere; keep stuff at -80 vs -70

# What can incentivize you to create technical notes?

- ☐ PI tells me to do it
- ☐ Special session at conference
- ☐ Grant/fellowship/award from some funding agency
- ☐ Published somewhere associated with “big name” journal
- ☐ \_\_\_\_\_?

# Bits and pieces

- How Do Astronomers Share Data? Reliability and Persistence of Datasets Linked in AAS Publications and a Qualitative Study of Data Practices among US Astronomers (2014)



Half-life of self-hosted content is ~10 years

