

How I Fail in Open Science

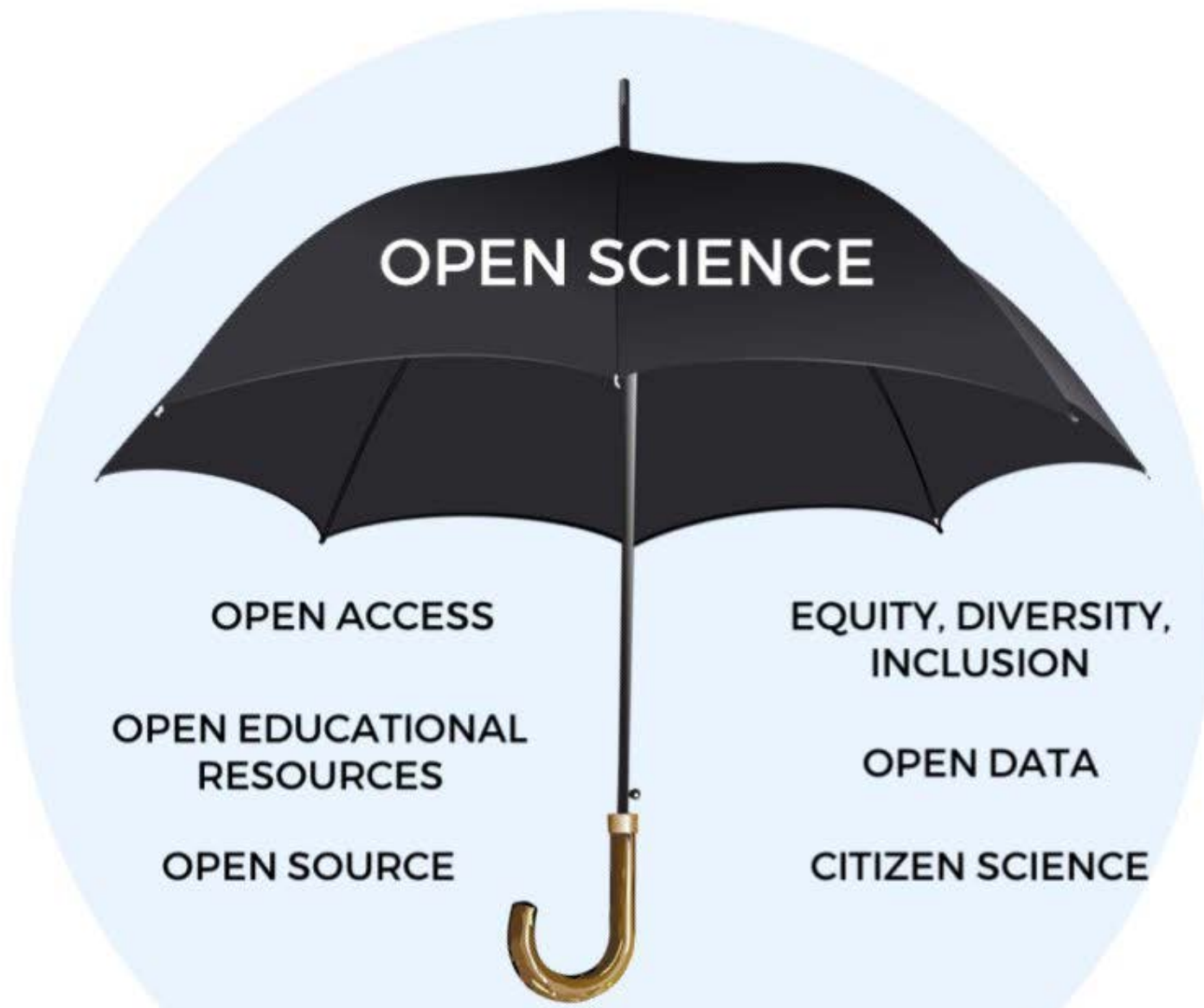
Veronika
Cheplygina

@drveronikach



<http://www.veronikach.com>

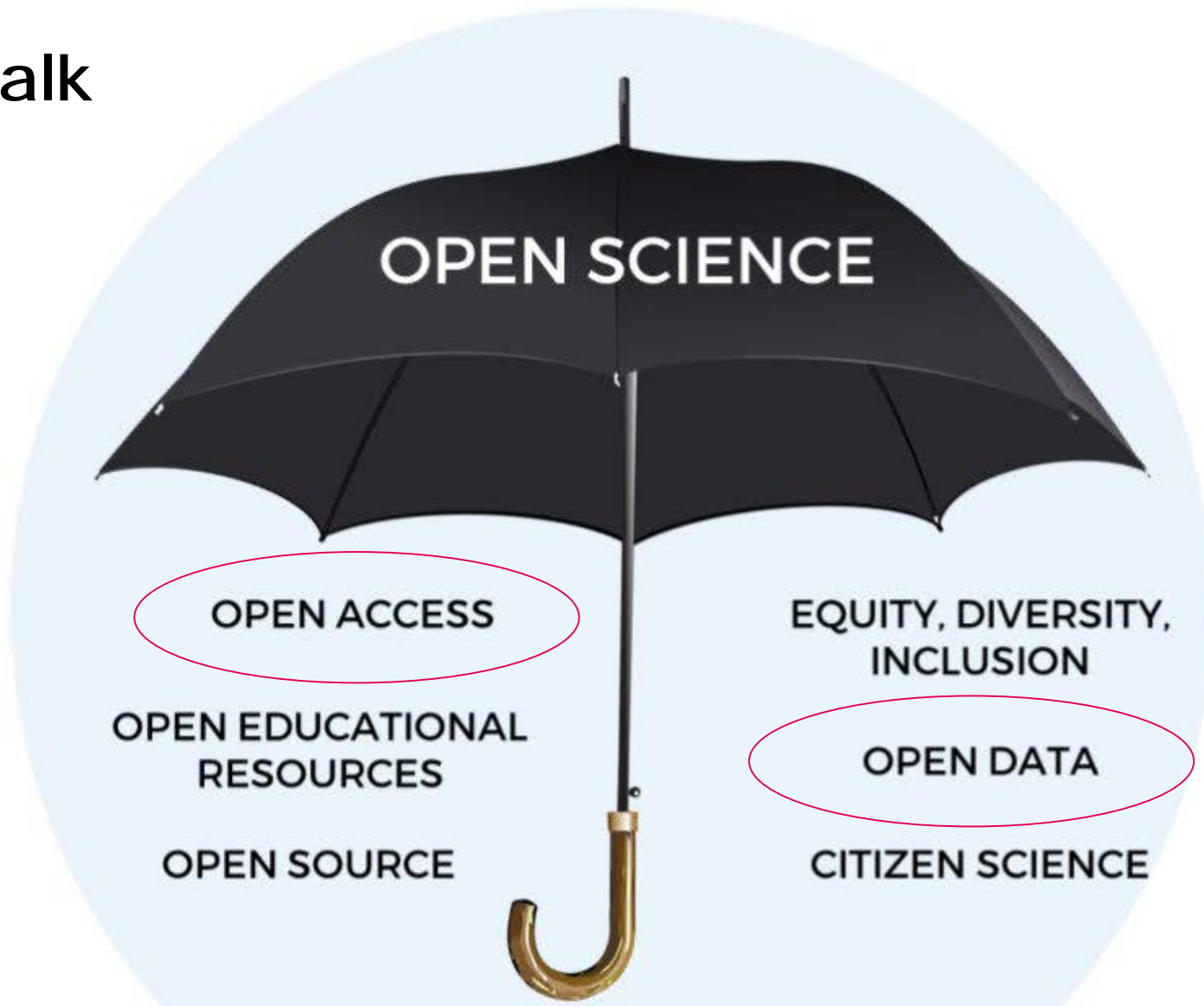




What do we mean when we talk about Open Science?

Image courtesy of Robin Champieux

This talk



What do we mean when we talk about Open Science?

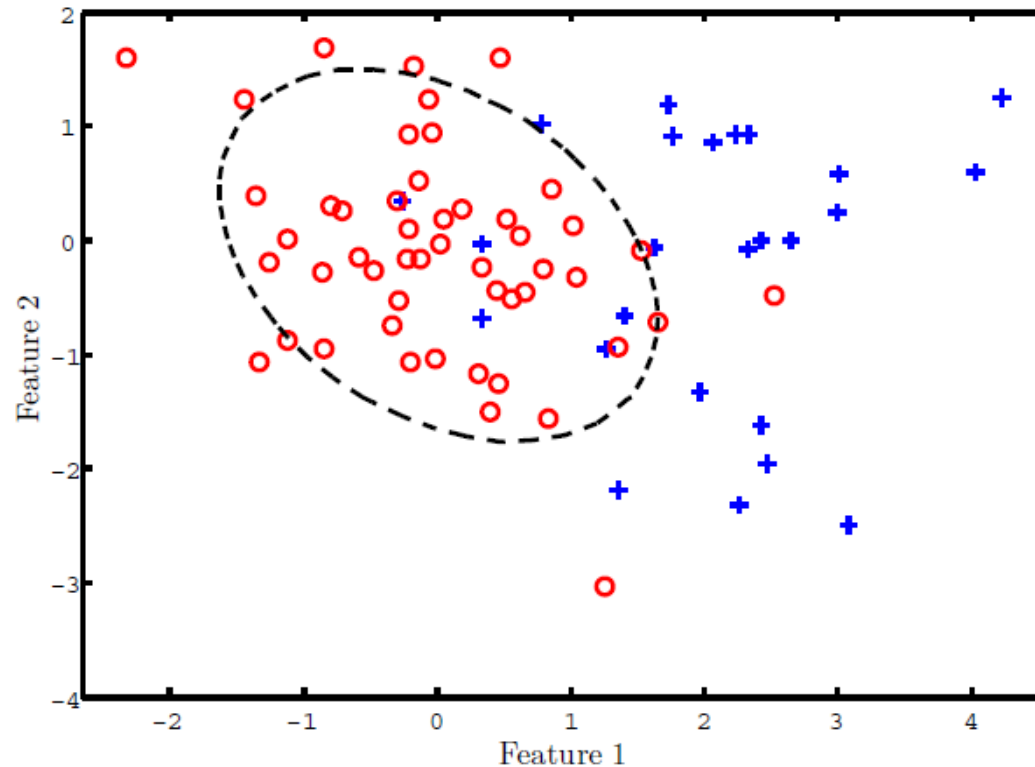
Image courtesy of Robin Champieux

This talk

- My experiences with open data & open access
- What did I (not) do & why?
- Some strategies to improve

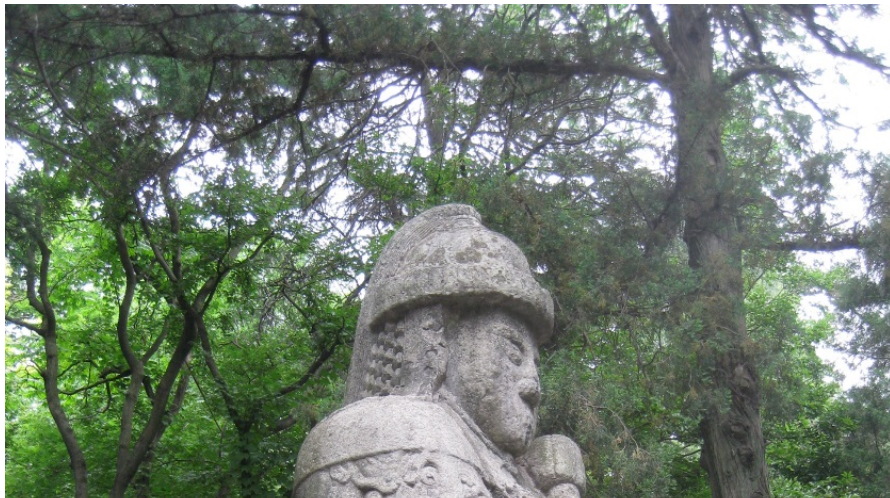
2011: Start PhD, publish papers!

Public clean data, own code
→ easy to share?

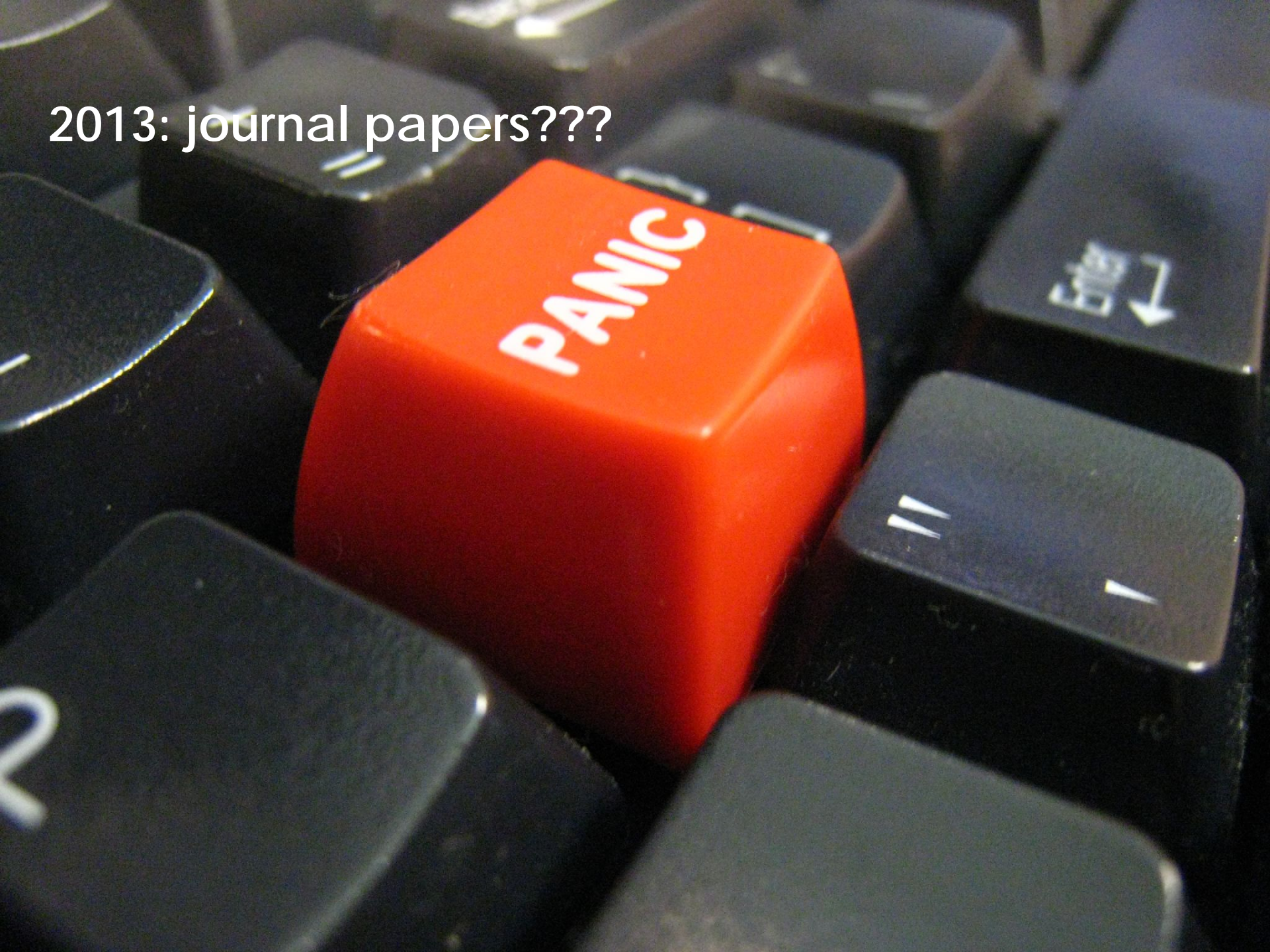




2011 - 2013: Workshop papers, no open science



2013: journal papers???



2013: writing journal papers

Publishing takes long → preprints!



SHERPA/RoMEO

Search - Publisher copyright policies & self-archiving

One journal found when searched for: 0278-0062

Journal:	IEEE Transactions on Medical Imaging (ISSN: 0278-0062)
RoMEO:	This is a RoMEO green journal
Paid OA:	A paid open access option is available for this journal.
Author's Pre-print:	✓ author can archive pre-print (ie pre-refereeing)
Author's Post-print:	✓ author can archive post-print (ie final draft post-refereeing)
Publisher's Version/PDF:	✗ author cannot archive publisher's version/PDF

<http://www.sherpa.ac.uk/romeo/search.php>

2013: writing journal papers

Procrastinating →
share data & code !

The CRAPL: An academic-strength open source license

[\[article index\]](#) [\[email me\]](#) [\[@mattmight\]](#) [\[rss\]](#)

Academics rarely release code, but I hope a license can encourage them.

Generally, academic software is stapled together on a tight deadline; an expert user has to coerce it into running; and it's not pretty code. Academic code is about "proof of concept." These rough edges make academics reluctant to release their software. But, that doesn't mean they shouldn't.

Most open source licenses (1) require source and modifications to be shared with binaries, and (2) absolve authors of legal liability.

An open source license for academics has additional needs: (1) it should require that source and modifications used to validate scientific claims be released with those claims; and (2) *more importantly*, it should absolve authors of shame, embarrassment and ridicule for ugly code.

Openness should also hinge on publication: once a paper is accepted, the license should force the release of modifications. During peer review, the license should enable the confidential disclosure of modifications to peer reviewers. If the paper is rejected, the modifications should remain closed to protect the authors' right to priority.

Toward these ends, I've drafted the **CRAPL--the Community Research and Academic Programming License**. The CRAPL is an open source "license" for academics that encourages code-sharing, regardless of how much how much Red Bull and coffee went into its production. (The text of the CRAPL is in the article body.)

<http://matt.might.net/articles/crapl/>

2015: journal papers!

Preprint bump



2015: postdoc, publish ASAP!

Closed data

Used existing code setup (suboptimal)

2016: find next (academic?) job → update CV

Blogging & Twitter



2017: start tenure track, do all the things!

Many responsibilities, failed to stay on top of goals



Still a lot to do

For 4 papers only 2x data, 2x code (0x works out of the box)

2018

Pena, Isabel Pino, Veronika Cheplygina, Sofia Paschaloudi, Morten Vuust, Jesper Carl, Ulla Møller Weinreich, Lasse Riis Østergaard, and Marleen de Bruijne. “Automatic Emphysema Detection using Weakly Labeled HRCT Lung Images.” PLoS ONE, 13(10): e0205397, 2018. [arXiv](#) | [Publisher](#) | [Data](#) | [Code](#)

Cheplygina, Veronika and Josien P W Pluim. “Crowd disagreement about medical images is informative”.

Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (MICCAI LABELS), pp. 105-111. Springer, 2018. [arXiv](#) | [Data](#)

Ørting, Silas, Jens Petersen, Veronika Cheplygina, Laura H. Thomsen, Mathilde M W Wille, and Marleen de Bruijne. Feature learning based on visual similarity triplets in medical image analysis: A case study of emphysema in chest CT scans.

Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (MICCAI LABELS), pp. 140-149. Springer, 2018. [arXiv](#)

Carbonneau, Marc-André, Veronika Cheplygina, Eric Granger, and Ghyslain Gagnon. “Multiple instance learning: A survey of problem characteristics and applications.” *Pattern Recognition* (2018). [arXiv](#) | [Publisher](#) | [Code](#)

Reviewer #2




“Need more high IF publications”

“THAT’s not how to do open science!”


Strategies
to do better




1. Start slow & focus on process


 habitica


TasksInventoryShopsPartyGuildsGroupChallengesHelp








Veronika
@veronikach • Level 80 Rogue

 36 / 50

 1392 / 2540

 190 / 190




Search


Tags


AllWeakStrong

Add a Habit


 Take stairs

» +0




 Do push-ups 🔥

» +0





Dailies1

Add a Daily


 Go for a run

» 0




 Walk 10K steps 🔥

» 1



To-Dos1

Add a To-Do

 Sign up for CPC 2019

2. Find accountability & support



3. Reward yourself




1. Start slow & focus on process



OpenMR talk




✓ Before talk:




- ☒ Send title, abstract, bio to host 26 Oct 2018
- ☒ Prepare slides OpenMR 2 Jan 18:00
- ☒ Announce on social media 2 Jan
- ☐ Update slides Today 09:00
- ☐ Practice talk  Today 17:00







✓ After talk:



- ☐ Add talk to CV Thursday
- ☐ Upload slides to website and/or repository Thursday
- ☐ Share slides on social media Thursday


2. Accountability & support




 WhitakerLab / WhitakerLabProjectManagement







 Watch 5  Star 13  Fork 17

 Code  Issues 90  Pull requests 1  Projects 1  Wiki  Insights


Branch: master ▾ WhitakerLabProjectManagement / README.md  Find file  Copy path

 stephwright Fixed typo ebcf6ac on Jun 2, 2017

3 contributors   

30 lines (15 sloc) | 2.31 KB  Raw  Blame  History   

Whitaker Lab Project Management

chat on  gitter

This repository contains issues related to the management of all Whitaker Lab projects along with weekly progress records for all lab members and collaborators.

Check out the [roadmap](#) for more information.

Working open

At the Whitaker lab we are dedicated to working as openly as possible. We believe that our endeavors will be more successful through clear communication and collaboration than by keeping secrets. This repository is designed to showcase our "work in progress". You're welcome to help out along the way!

<https://github.com/WhitakerLab/WhitakerLabProjectManagement>

2. Accountability & support

#tips

☆ | 👤 8 | 📌 0 | ✎ Add a topic



December 30th, 2018



veronika 12:14 PM

how/where would you do that? what I'm trying now in my todo list (in Todoist) is that I have a list of current projects, then one for 2019 - quarter 1 (based on deadlines and such), and then "Later". This is how I can see that I already planned too much for the first quarter 😊

kind of similar to Kanban idea, but larger scale



mdekstrand 🐰 3:04 PM

probably paper

Wednesday, January 2nd



noeska 12:58 PM

Hmm, I like this paper calendar idea, maybe something to hang in my office 😊

Friday, January 4th



felienne 1:32 PM

We have a timeline for the whole year in our office with deadlines for conferences on which people can also add other goals, works pretty nicely (also to hold each other accountable)



veronika 1:43 PM

Cool!

Tuesday, January 8th

2. Accountability & support



Spread the Word

Open Science is for everyone. Help us spread the word by [joining our open Slack channel](#) and invite your friends and colleagues to be part of our community.



Join Our Team

Everything we develop and produce is in the open. You can [join our team on GitHub](#). If you want to become an instructor [email us](#).

<https://opensciencemooc.eu/>

3. Reward yourself



7 ACHIEVEMENTS



Open Access Top 10%

90% of your research is free to read online. This level of availability puts you in the top 2% of researchers.

[link](#) [share](#)



Wikistatic Top 25%

Your research is mentioned in 2 Wikipedia articles! Only 20% of researchers are this highly cited in Wikipedia.

[Your Wikipedia titles include Random subspace method., Multiple instance learning](#)

[link](#) [share](#)



All Readers Welcome Top 25%

Your writing has a reading level that is easily understood at grade 13 and above, based on its abstracts and titles.

[link](#) [share](#)



Global Reach

Your research has been saved and shared in 18 countries.

[Countries include Australia, Brazil, Canada and 15 more.](#)

[link](#) [share](#)



Greatest Hit Top 50%

Your top publication has been saved and shared 30 times. Only 29% of researchers get this much attention on a publication.

[Your greatest hit online is Transfer learning for multi-center classification of chronic obstructive pulmonary disease.](#)

[link](#) [share](#)

3. Reward yourself



Mon, Nov 19, 2018, 2:22 AM



We just covered your arXiv paper “**Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis,**” in my Medical Imaging and Machine Learning course and the students really liked reading it. Great job and congratulations – it must have been a lot of work - it makes for a great resource for keeping up with research as well as for teaching.

Best Regards,



How (not to?) Fail in Open Science

- Start slow & focus on process
- Find accountability & support
- Reward yourself!

?

T₁

H₄

A₁

N₁

K₅

S₁

@drveronikach



<http://www.veronikach.com>

