GitLab as an Institutional Service
Conquaire workshop, 3 April 2019, Christian Pietsch

At the end of this talk, you will …

- understand why versioning is important
- know what Git is
- know what GitHub is
- know what GitLab is
- know what Continuous Integration is.

Daily routine of knowledge workers:

(1) from one island of stability
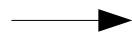(2) through uncertain territory
(3) to another island.

- *Undo*

- *Record changes*

- *Save* in a wiki

- *Transparent version tracking* in some file systems, online storage

- *Backups*

- Inventive use of *file names*:
  - cv-old.doc
  - cv-old-old.doc
  - cv-2019.doc
  - cv_2019.doc
  - cv-april.doc
  - cv-2019-backup.doc
  - cv-final.doc
  - cv-really-final.doc

Why?

Versions are the
stepping stones of
knowledge workers.

(1) island of stability ➝
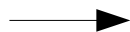
(2) uncertain territory ⟶

(1) island of stability ⟶

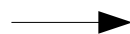(3) new island →

(2) uncertain territory →

(1) island of stability →
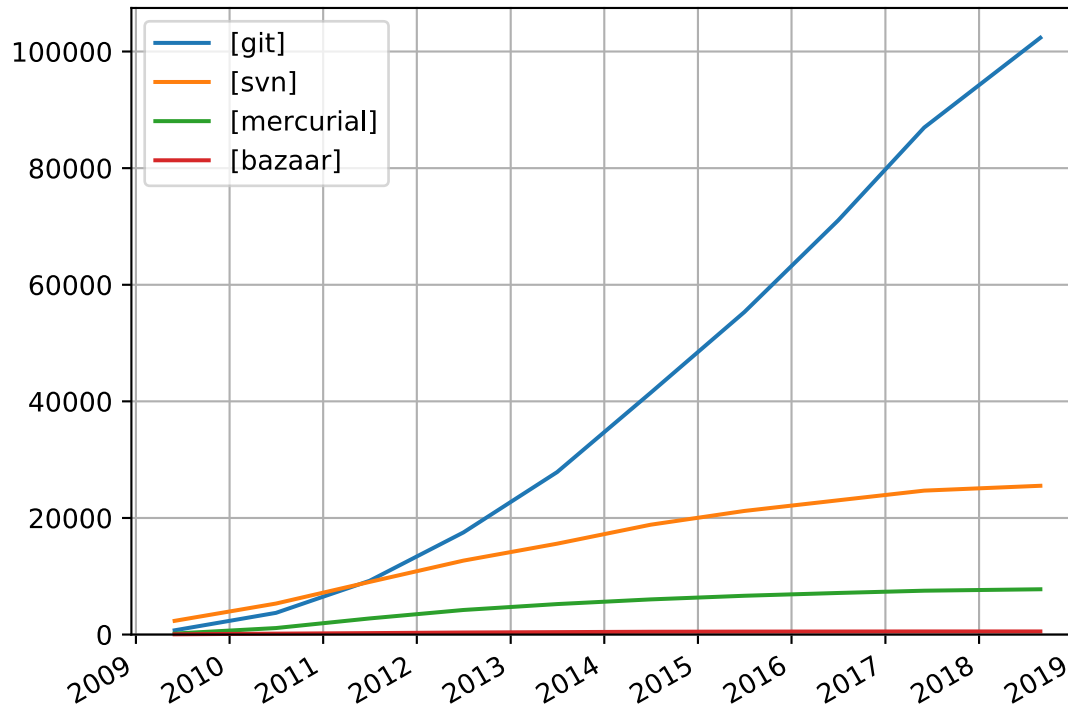
(3) new island →

(2) uncertain territory →

(1) island of stability →

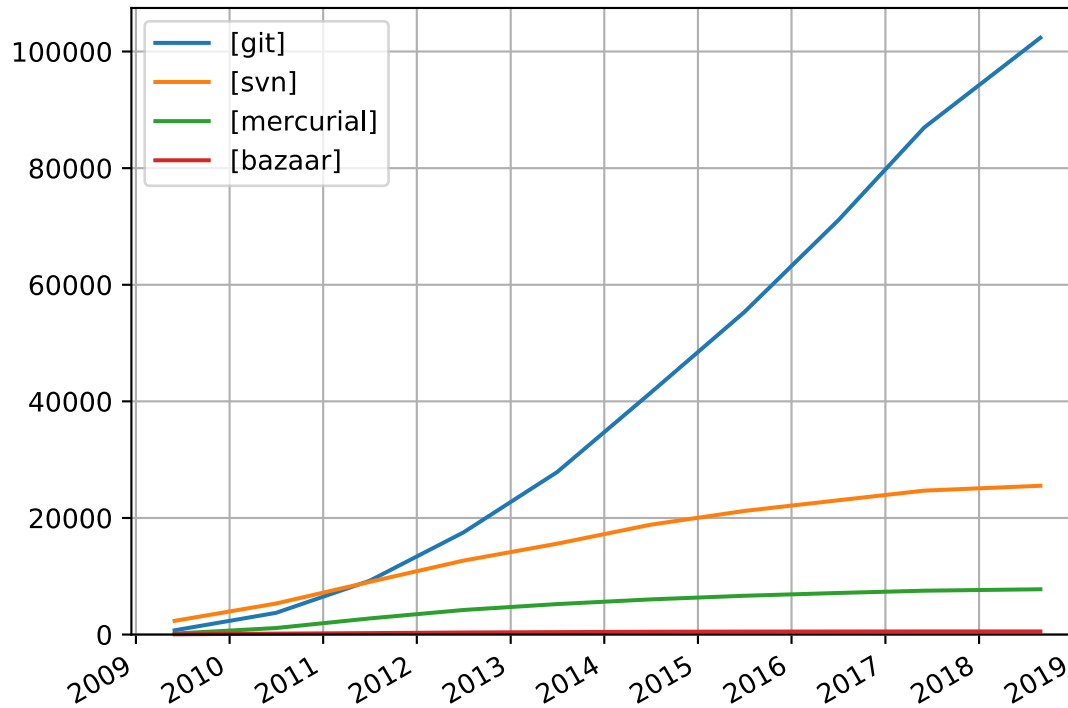# Popular versioning software

number of questions asked on StackExchange about current versioning software

Universität Bielefeld
INFORMATION.*plus!*
Universitätsbibliothek

GitHub's popularity boosts Git

2008: GitHub launch

2010: GitHub hosting 1 million repositories

2011: GitHub hosting 2 million repositories, surpassing SourceForge and Google Code

2013, January: GitHub hosting 5 million repositories

2013, December: GitHub hosting 10 million repositories

Why is GitHub so popular?
Social coding!

1. You publish your code in your own GitHub repository

2. Others "fork" (copy) your code and improve it in their own GitHub repository

3. Others give back their improvements via "pull requests"

4. You decide whether you accept their code donation

5. You benefit without giving up control

What is the problem with GitHub?
Vendor lock-in!

- GitHub is not free software

- On-premises hosting is possible, but you are still depending on GitHub

- Single point of failure (largest DDOS attack ever)

- GitHub may go bankrupt or be acquired

- GitHub was acquired by Microsoft in 2018
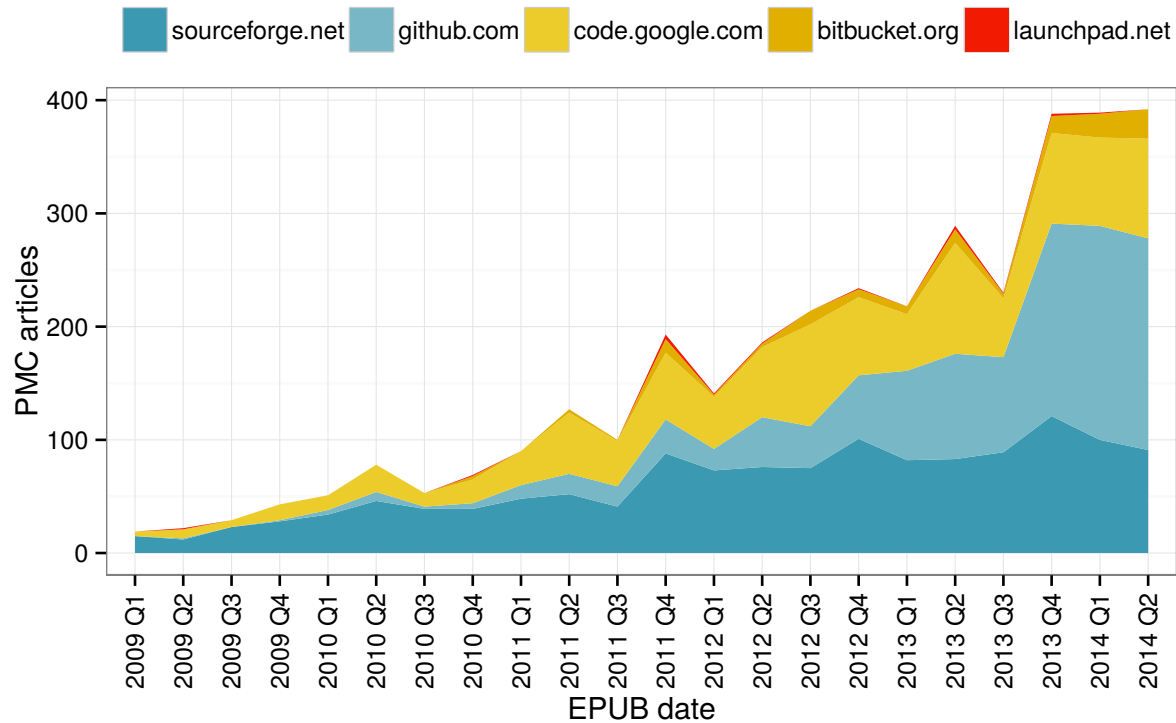
Are there alternatives to GitHub?
Yes!

- GitLab CE is free software

- GitLab CE includes enterprise features such as SSO

- On-premises hosting is easy – one big Linux package

- If GitLab is acquired by, say, Google, the GitLab code could be forked and developed independently
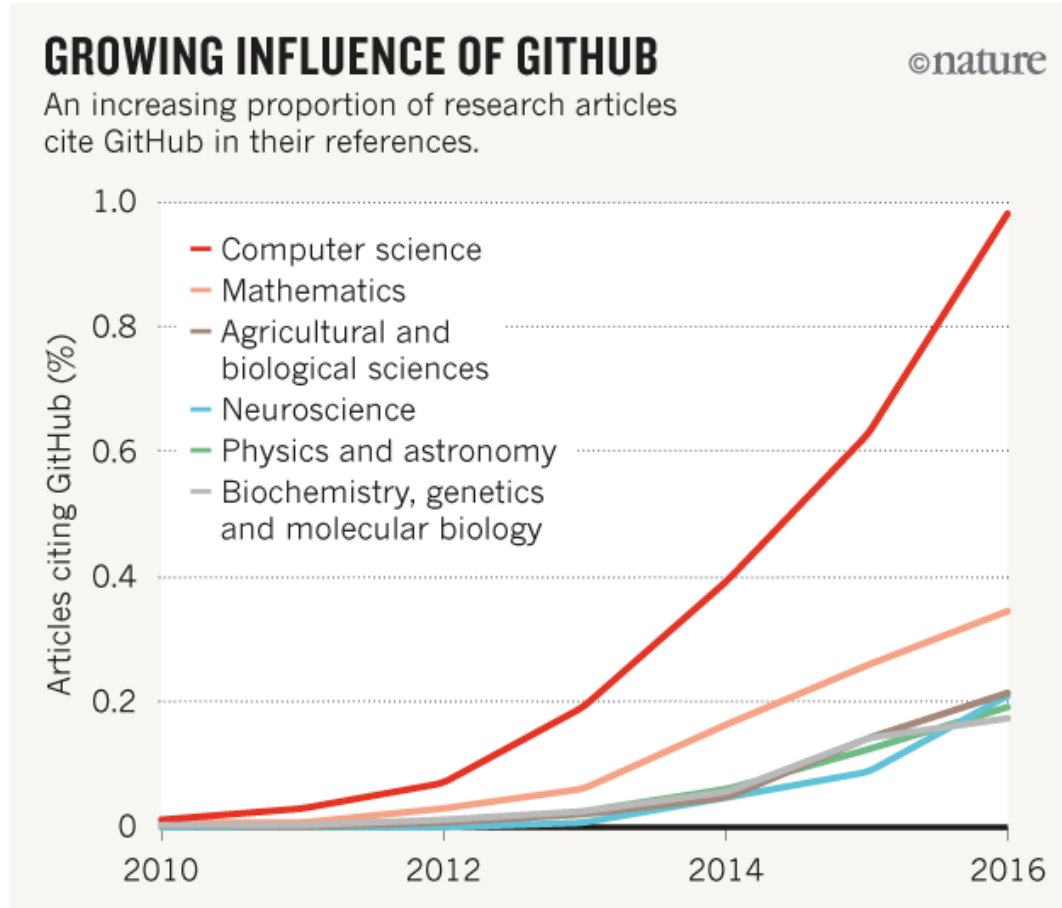
- Let's try it: https://gitlab.ub.uni-bielefeld.de

Researchers discover GitHub [Ram, 2013]

"While Git provides powerful features that can leveraged by individual scientists, Git hosting services open up a whole new set of possibilities. Any local Git repository can be linked to one or more Git remotes, which are copies hosted on remote cloud servers. Git remotes serve as hubs for collaboration where authors with write privileges can contribute anytime while others can download up-to-date versions or submit revisions with author approval. There are currently several Git hosting services such as SourceForge, Google Code, GitHub, and BitBucket that provide free Git hosting. Among them, GitHub has surpassed other source code hosts like Google Code and SourceForge in popularity and hosts over 4.6 million repositories from 2.8 million users as of December 2012."

Karthik Ram (2013). Git can facilitate greater reproducibility and increased transparency in science. Source Code for Biology and Medicine, 8:7. doi: 10.1186/1751-0473-8-7

Quarterly distribution of links to open source hosting facilities in Europe PMC full text articles (research by Najko Jahn published in the Conquaire project proposal. doi:10.5281/zenodo.31298)

Jeffrey Perkel (2016). Democratic databases: science on GitHub.
In: Nature 538:7623, 6, p. 127–128, doi:10.1038/538127a

Conquaire = GitLab + Continuous Integration

made usable for researchers of all areas, locally

Things to remember

- Editing files without proper versioning is like aerial acrobatics without a safety net.

- Git is the versioning software to use. Yes you can.

- GitHub and GitLab are great for collaboration.

- Hosting GitLab on premises is free and easy.

- Continuous Integration can save your day.