Computer Tools for Efficient Science

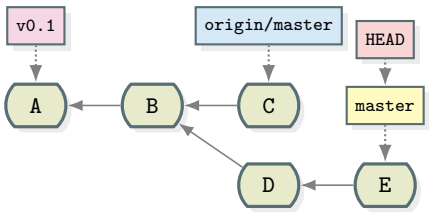
Daan van Vugt daanvanvugt@gmail.com

2/12/2015

Version control

Git

- Git manages changes to a tree of files over time
- Distributed development, many branches
- Excellent integration with many sites and services (GitHub, GitLab, Bitbucket)



See also: Mercurial (hg), svn

Github

> GitHub is how people build software. With a community of more than 11 million people, developers can discover, use, and contribute to over 29 million projects using a powerful collaborative development workflow. (source: github.com/about)

Alternatives

- Bitbucket (with free private projects)
- GitLab (self-hosted)

Writing papers/reports

Use LaTeX

Why?

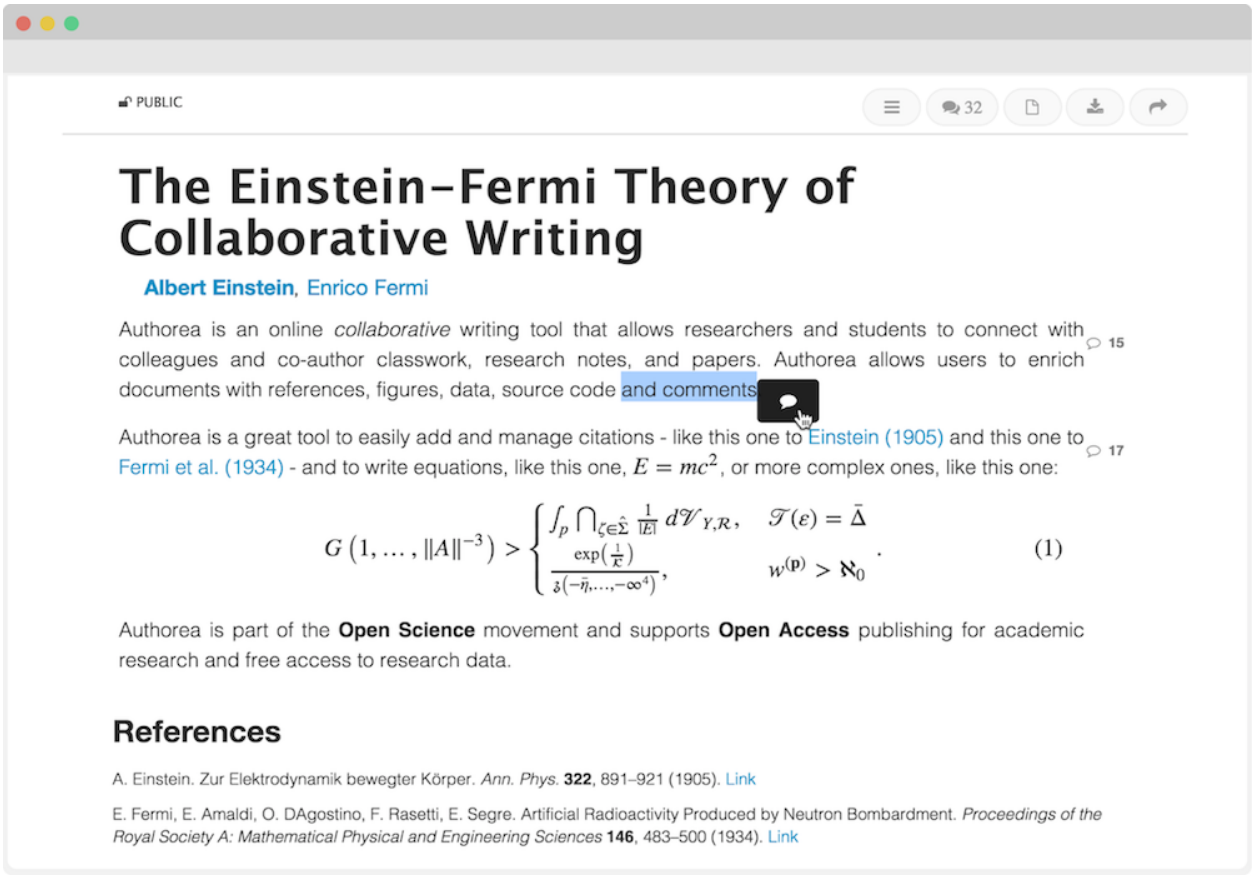
- Easy collaboration
- Version control

- Bibtex integration
- Formulas are easy
- It looks awesome
- Automation

How?

- Web-based: **Authorea**, ShareLaTeX, WriteLaTeX, StackEdit
- Windows/Mac: TeXMaker/TeXstudio/TeXshop
- Linux: Gummi, Vim, Emacs (+Auctex)

Authorea



Alternatives to LaTeX

- Writefull (checks text for correct language)
- Draft & Typewrite (Real-time collaborative writing)
- Hackpad
- Etherpad
- Google Drive
- Microsoft Word (eww)

Sharing data accompanying the paper

- 3TU datacenter
- DataCite
- Dryad
- Figshare
- OpenScienceFramework
- Slideshare
- Zenodo

Interactive notebooks

Write your code and documentation in the same place

- Jupyter (IPython/Ruby/Julia, link to Authorea, SageMathCloud)
- Mathematica
- rCharts + Slidify + shiny (R)

Writing code

Best practices

- **Version control!**
- Unit tests
- Documentation
- Read about best practices online
- Version control (again, because it's important)
- Check the style guide for your language/project
- Did I mention version control?

Editors

- Vim (vimtutor & vim-adventures to learn)

- Emacs
- Notepad++
- Sublime Text
- Many more, choose one you like and pimp it, add syntax highlighting etc

Many offer integration with syntax checkers and build tools

(Hard mode) Learn a better keyboard layout: (Programmer) Dvorak or Colemak

Analyzing data

MATLAB

Good features

- Contains a unit test framework since 2013
- Some integration with git
- Nice GUI, tools like profiler and parfor

Open alternatives:

- Octave
- Python + Numpy + Scipy
- Linux tools: sed, awk, grep, gnuplot etc.
- R, paraview
- C/C++/Fortran shlib + python (for speed)

Creating figures

Tools for generating graphics

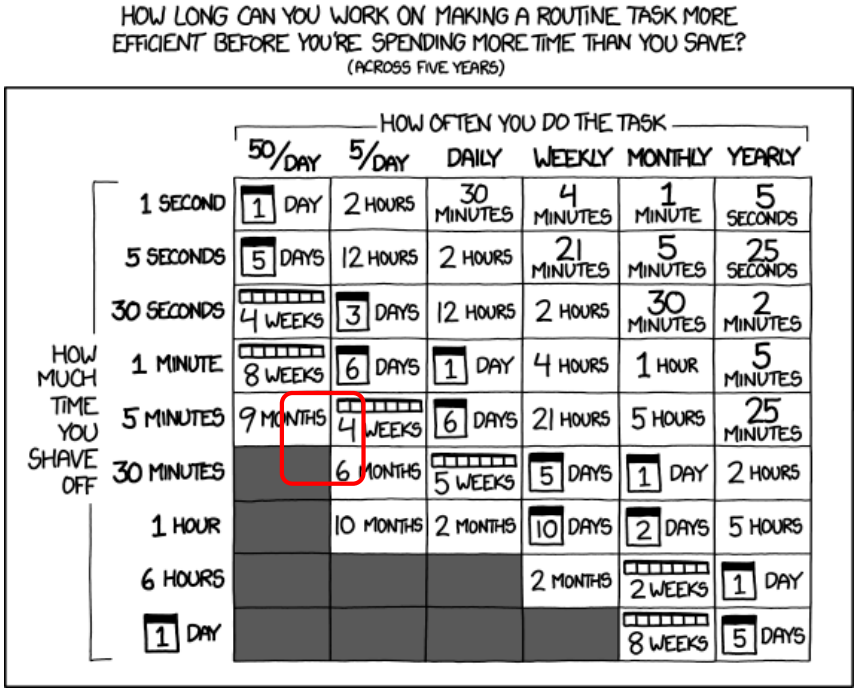
- D3.js (Interactive on webpages)
- Matplotlib (Python, + D3.js)
- Gnuplot
- MATLAB / Octave + Matlab2tikz (for LaTeX)
- Ggplot2 (R)
- Mathematica / Maple
- Paraview (3D figures)

Why not Excel/Origin?

- Not easily scriptable / automated
- Hard to create publication-quality graphics

Automation

Is It Worth The Time? (XKCD)



How to automate

- GNU Make
- Bash scripting
- Ruffus (Python, computational pipelines)
- Snakemake, Pegasus (Workflow management system)
- Vistrails (Workflow and provenance management system)
- PyRDM (Research Data Management)
- Sumatra, Elabftw, Wings Workflow (Electronic Lab Notebook)
- Digital lab notebooks: Evernote, Onenote

Snakemake

Presenting research

Posters

- LaTeX (poster package)
- Scribus
- Inkscape
- Powerpoint

Presentations

- LaTeX Beamer class (this presentation, + Pandoc)
- Prezi
- Powerpoint

Searching / reading papers

Mendeley/Zotero/JabRef/Bibdesk

Mendeley:

- Import papers, automatically gets name and title right
- Share libraries with colleagues
- Sync bibtex files with LaTeX
- Full-text search

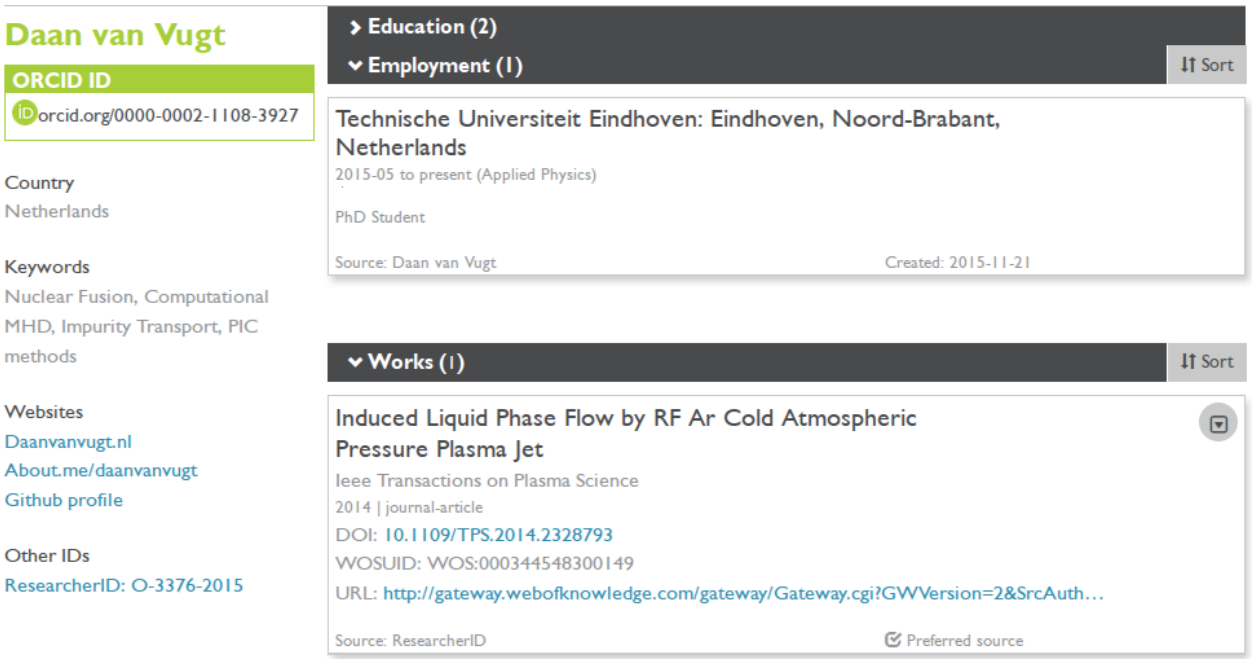
Other tools, sites for finding papers

- Web of Science
- Google scholar
- Webplotdigitizer
- CiteULike
- ResearchGate
- Scopus
- Lazyscholar.org

Online presence

- ORCID
- ResearcherID
- Academia.edu
- ResearchGate
- About.me
- Twitter

ORCID



Now it is your turn!

- Read 5-10 minutes about some of these programs
- Try one (or a few)
- Let me know how it goes, and if you find something interesting
- View the source of this presentation on https://github.com/Exteris/tools-for-science