

Computer Tools for Efficient Science

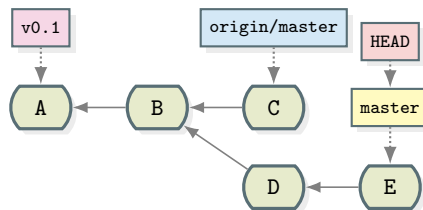
Daan van Vugt daanvanvugt@gmail.com

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

The screenshot shows a web browser window displaying an Authorea document. The document title is "The Einstein-Fermi Theory of Collaborative Writing" by Albert Einstein and Enrico Fermi. The text describes Authorea as an online collaborative writing tool. It includes a paragraph about adding citations and writing equations, followed by a complex mathematical equation labeled (1). The document also has a "References" section with two entries: A. Einstein (1905) and E. Fermi et al. (1934). The interface includes a "PUBLIC" status indicator, a menu icon, and icons for comments (32), document, download, and share.

Public

The Einstein-Fermi Theory of Collaborative Writing

Albert Einstein, Enrico Fermi

Authorea is an online *collaborative* writing tool that allows researchers and students to connect with colleagues and co-author classwork, research notes, and papers. Authorea allows users to enrich documents with references, figures, data, source code and comments.

Authorea is a great tool to easily add and manage citations - like this one to Einstein (1905) and this one to Fermi et al. (1934) - and to write equations, like this one, $E = mc^2$, or more complex ones, like this one:

$$G(1, \dots, \|A\|^{-3}) > \begin{cases} \int_p \bigcap_{\zeta \in \hat{\Sigma}} \frac{1}{|\bar{E}|} d\mathcal{V}_{Y, \mathcal{R}}, & \mathcal{T}(\varepsilon) = \bar{\Delta} \\ \frac{\exp(\frac{1}{\varepsilon})}{\delta(-\bar{\eta}, \dots, -\infty^4)}, & w^{(p)} > \aleph_0 \end{cases} \quad (1)$$

Authorea is part of the **Open Science** movement and supports **Open Access** publishing for academic research and free access to research data.

References

A. Einstein. Zur Elektrodynamik bewegter Körper. *Ann. Phys.* **322**, 891–921 (1905). [Link](#)

E. Fermi, E. Amaldi, O. D'Agostino, F. Rasetti, E. Segre. Artificial Radioactivity Produced by Neutron Bombardment. *Proceedings of the Royal Society A: Mathematical Physical and Engineering Sciences* **146**, 483–500 (1934). [Link](#)

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)

```
In [9]: display(i)
```

IP[y]: IPython
Interactive Computing

```
In [3]: from IPython.display import SVG  
SVG(filename='python-logo.svg')
```

Out[3]:

The image shows a Jupyter notebook interface. It contains two code cells. The first cell has the code 'display(i)' and its output is 'IP[y]: IPython Interactive Computing'. The second cell has the code 'from IPython.display import SVG' and 'SVG(filename='python-logo.svg')', and its output is the Python logo, which consists of two interlocking snakes (one blue, one yellow) and the word 'python' in a grey sans-serif font with a trademark symbol.

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)

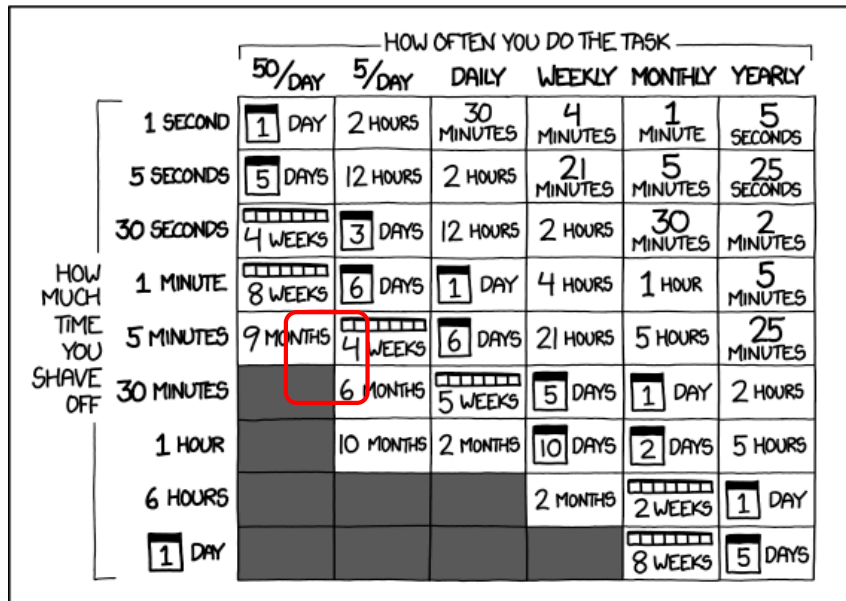
Why not Excel/Origin?

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

Automation

Is It Worth The Time? (XKCD)

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE?
(ACROSS FIVE YEARS)



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

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

Daan van Vugt

ORCID ID

 orcid.org/0000-0002-1108-3927

Country
Netherlands

Keywords
Nuclear Fusion, Computational MHD, Impurity Transport, PIC methods

Websites
Daanvanvugt.nl
About.me/daanvanvugt
[Github profile](https://github.com/daanvanvugt)

Other IDs
ResearcherID: O-3376-2015


Education (2)

Employment (1) Sort

Technische Universiteit Eindhoven: Eindhoven, Noord-Brabant, Netherlands
2015-05 to present (Applied Physics)
PhD Student
Source: Daan van Vugt Created: 2015-11-21

Works (1) Sort

Induced Liquid Phase Flow by RF Ar Cold Atmospheric Pressure Plasma Jet 

Ieee Transactions on Plasma Science
2014 | journal-article
DOI: [10.1109/TPS.2014.2328793](https://doi.org/10.1109/TPS.2014.2328793)
WOSUID: WOS:000344548300149
URL: <http://gateway.webofknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcAuth...>
Source: ResearcherID  Preferred source

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>