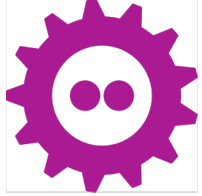
FOSDEM 2020 

fosdem.org

- Free and Open source Software Developers' European Meeting
- Brussels 1-2 February 2020
- 8000+ participants
- 837 talks
- 100% live streamed
- talks available online

Schedule

- 56 parallel developer rooms
- 7 main tracks
    - Community and Ethics
    - Containers and Security
    - Databases
    - Freedom
    - History
    - Internet
    - Miscellaneous

Parallel tracks (devrooms)

- Continuous Integration and Continuous Deployment
- Free tools and editors
- HPC, Big Data, and Data Science
- Open Document Editors
- Python
- Quantum Computing
- Open Research Tools and Technologies devroom **(first time!)**

Interesting talks

- Double your contributors using these 3 simple tricks! - Why would someone work on your project?
- Interactive applications on HPC systems - Jupyterhub, Galaxy, RStudio, XPRA
- Production-time Profiling for Python
- Will somebody *please* tell me what's going on? Managing change in Python projects
- Asyncio: understanding async and await in Python
- How to write a scikit-learn compatible estimator/transformer
- Correlation analysis in automated testing
- How to fail successfully and reliably - And look good while doing it

Open Research Tools and Technologies (devroom)

Open Research Tools and Technologies devroom

1. Open software development in *research*
2. Open software in the social sciences / humanities / journalism

- mainly tool focused talks
- conceptual overlap with humanities

Open Research Tools and Technologies (devroom)

- *Jan Grewe* The good and the bad sides of developing open source tools for neuroscience
  - advantages and disadvantages of using OSS in science

- *Julia Sprenger* Challenges and opportunities in scientific software development
  - differences between scientific and classical software development
  - actions required on all sides to improve situation

- *Aniket Pradhan* NeuroFeodora: Enableing Free/Open Neuroscience
  - an OS for neuroscience research
  - 130 neuroscience packages ready to use
  - 120+ more packages will be added

- *Jan Hanke* DataLad
  - general-purpose tool to manage data across different repos
  - capable of handling large datasets via git-annex
  - provenance tracking & complete decentralization

- *Lilly Winfree* Frictionless Data for Reproducible Research
  - metadata aggregation, validation and curation pipelines for oceanographic data
  - open knowledge foundation
  - frictionlessdata & data validation goodtables

Open Research Tools and Technologies (devroom)

- *Mateusz Kuzak* On the road to sustainable research software
    - 4OSS simple recommendations (OSS from the beginning)
    - Develop a publicly accessible open-source code from day one.
    - Discoverable software by providing software metadata via a popular community registry.
    - Adopt a license and comply with the licence of third-party dependencies.
    - Clear and transparent contribution, governance and communication processes.

- *Sébastien Rochette* Transforming scattered analyses into a documented, reproducible and shareable workflow
    - *collaboration fest*, R workshop for supporting software development in the sciences

- *Emmy Tsang* (Innovation Community Manager eLife) A community-driven approach towards open innovation for research communication
    - feedback is essential & always welcome
    - recent innovation example: *live paper*

- *Karthik Ram* The Journal of Open Source Software Credit for invisible work
    - paper publication for mature software packages in the sciences
    - minimal effort for deserved credit (`2h paper, 1-2pages`)

Check out fosdem.org!

Notes

Feedback to Challenges and opportunities in scientific software development - Challenges and opportunities in scientific software development - questions - Software carpentry - For general programming skills, not for our particular softwares - How to better connect scientists to not duplicate software? - NFDI - comments - startup idea for collaborative workspace platform based on Jupyter Notebooks - humanities have dedicated software engineers funded - only innovative business departments use agile methods (e.g. some VW departements) - how to convince research groups to test agile methods?

Highlights from other talks

open knowledge foundation

- making data available
- organize: licencing, data cleaning, ..
- dedicated data managers for cleaning, annotation of data, quality assurance, gaps...

create.frictionlessdata.io

- metadataformat: json
- packaging data + metadata
- validation: try.goodtables.io
- mistaken identifiers: gene name errors when using excel
- datapackage-pipelines
- bco-dmo.org

A community-driven approach towards open innovation for research communication - Emmy Tsang (Jounal of open source science)