Sunet Drive

An Academic Toolbox

FAIR Data Storage, Analysis and Publication

Sunet Drive – Architecture
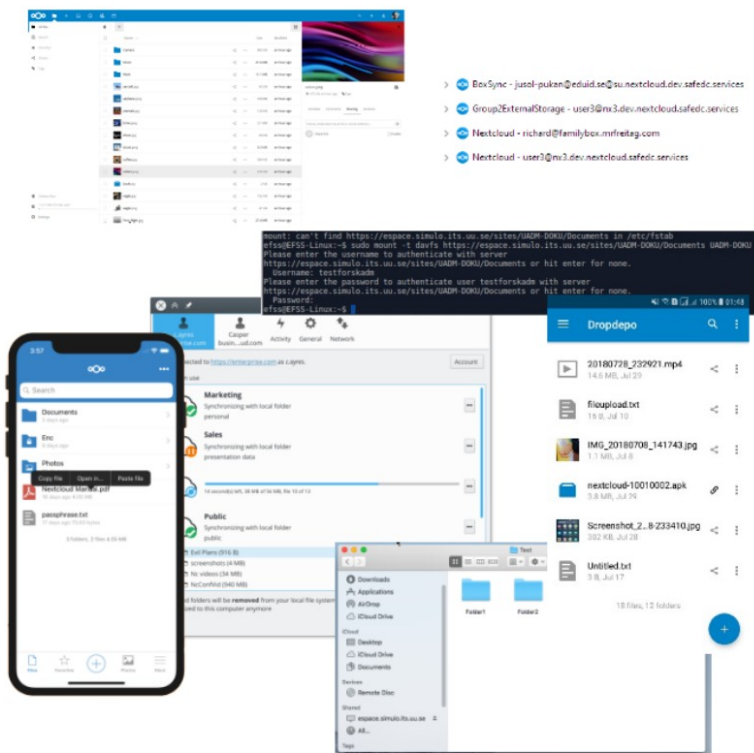
Primary data sources

Secure Storage

Analyze and Publish

UV-Vis

IVIUM

...

PIA

IPCE

Sunet Drive – Architecture

Primary data sources

Secure Storage

Analyze and Publish

UV-Vis

PIA

IPCE

An Academic Toolbox

Following FAIR principles
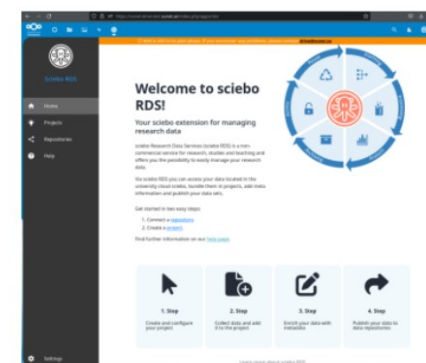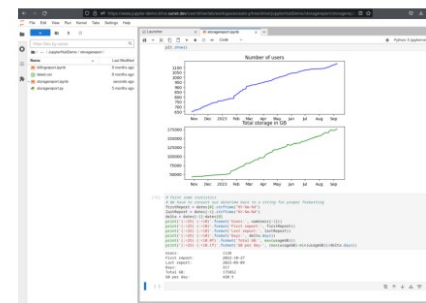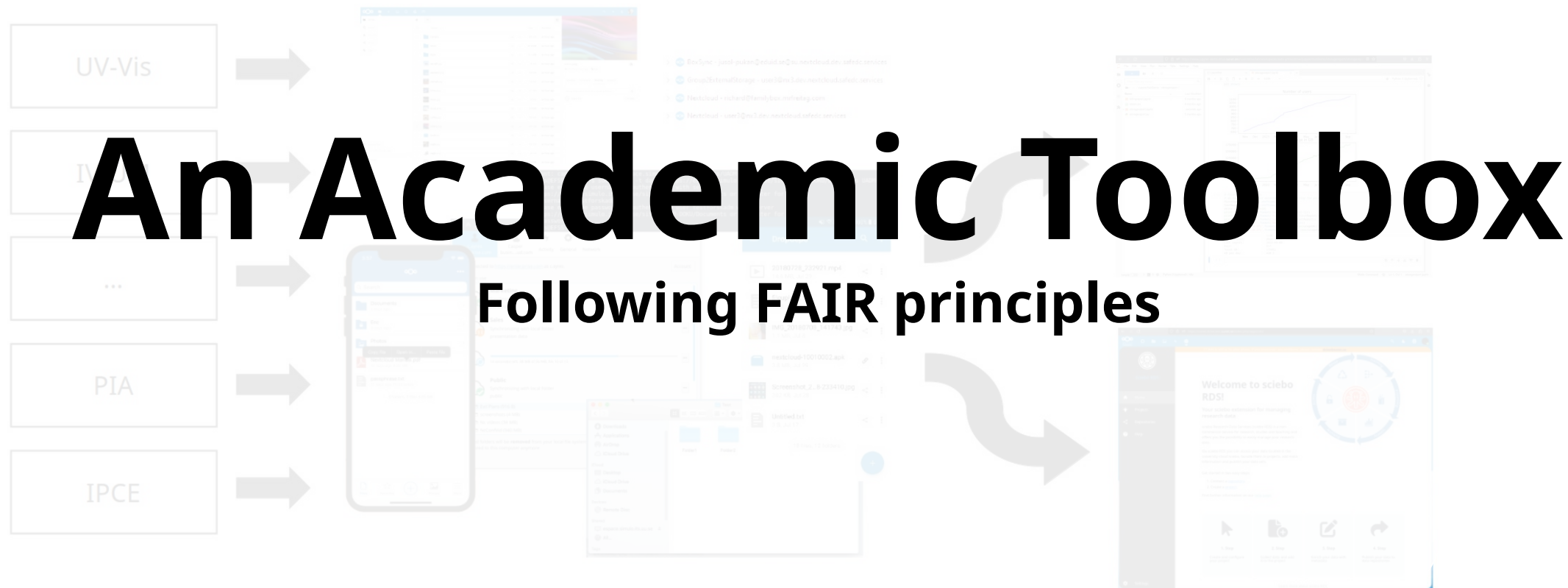
Sunet drive today

- 2x54 nodes on 300 servers in two data centres
- 100% powered by green energy
- 20 paying customers
- 508TB data stored with >0.5TB added daily
- 3,752 total users, >600 daily users

Number of users for su

Total storage in GB for su
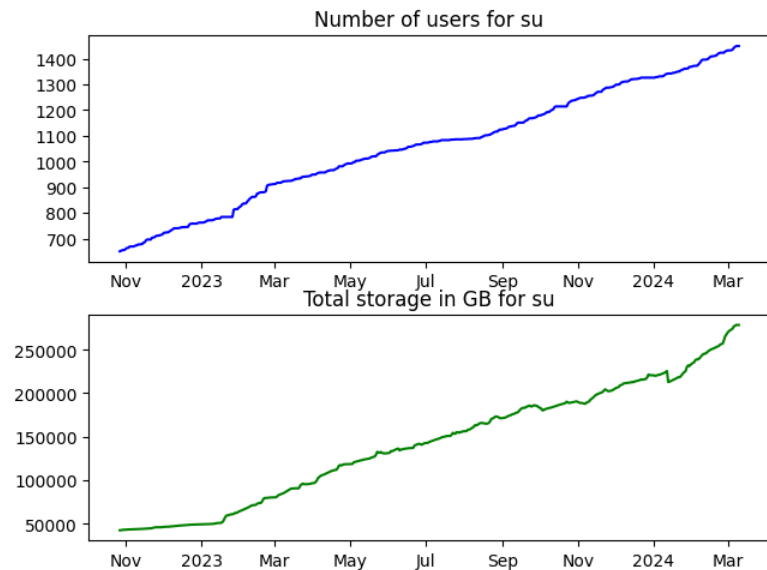
Sunet Drive development & projects

- Full RDS-setup in production
- JupyterHub
- Secure Zones/MFA/Step-up Authentication
- Security Audit
- Devops: Third data centre & Kubernetes
- AI Assistant with "Ethical AI"

Core Application
JupyterHub

Core Application
Sciebo Research Data Service

Core Application
ScienceMesh

SUNET

Sunet Drive

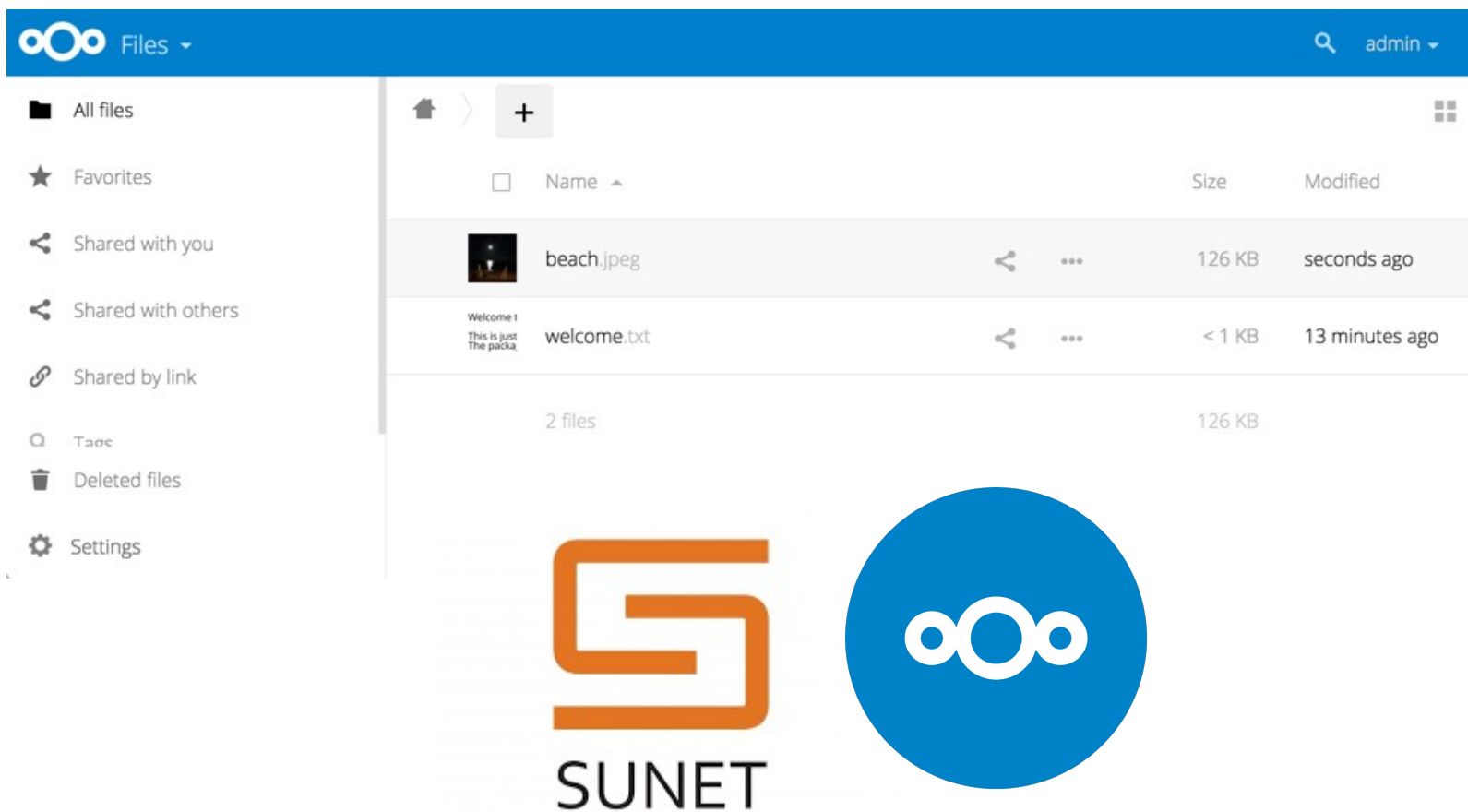- Allows people to store, share and collaborate on any type of file

- Runs in the browser, on your computer, on your phone and probably more!

- Secure datacenters controlled by SUNET. No data leaves our datacenter, unless you want it to!

Co-management of nodes

- S3-bucket storage

- Share data between nodes

- Access ownership separation

- Lifecycle managed buckets

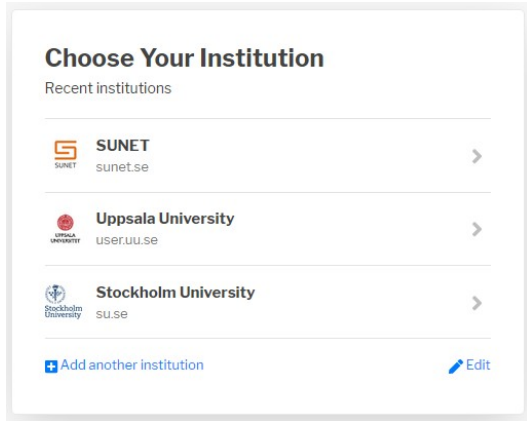- Co-managed with users for continuous improvement
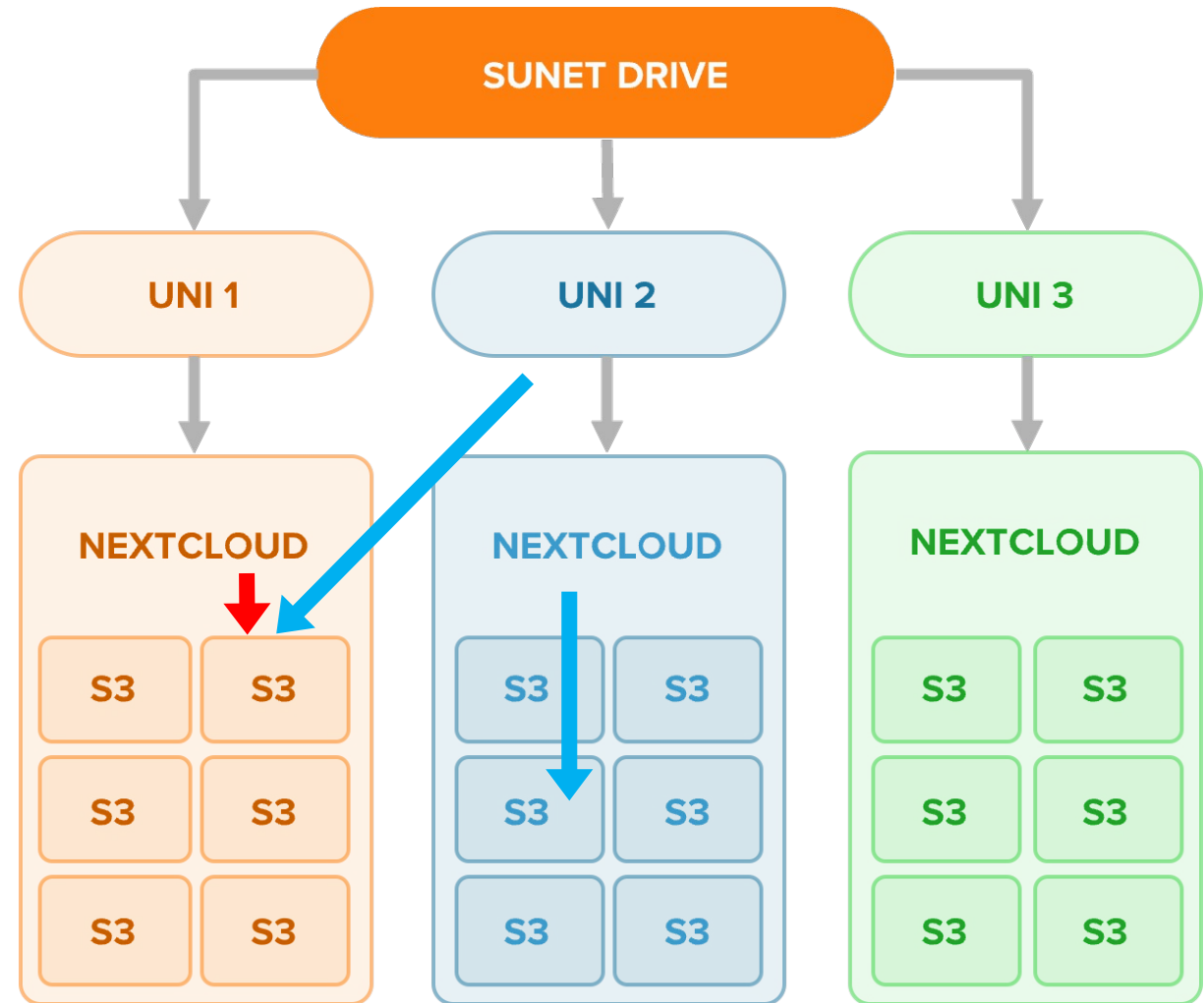
Global Collaboration

Nationally: eduGAIN

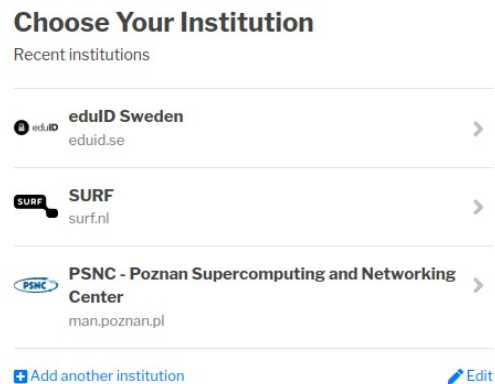- ~6000 Identity Providers
- Eduid.se for external collaborators

Choose Your Institution
Recent institutions

⬡ SUNET	sunet.se	>
🔴 Uppsala University	user.uu.se	>
Stockholm University	su.se	>

➕ Add another institution ✏ Edit

→ sunet.drive.sunet.se

→ uu.drive.sunet.se

→ su.drive.sunet.se

Choose Your Institution
Recent institutions

eduID Sweden	eduid.se	>
SURF	surf.nl	>
PSNC - Poznan Supercomputing and Networking Center	man.poznan.pl	>

➕ Add another institution ✏ Edit
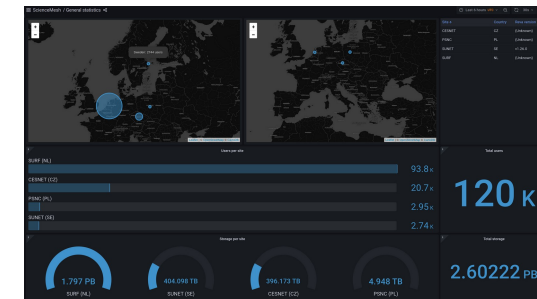
Swedish Account →

Global Account → → extern.drive.sunet.se

Globally: OCM

- OpenCloudMesh protocol
  - Initiated in 2015 by GÉANT
  - Supported by ownCloud, Nextcloud, Seafile
- Global federated shares with local sovereign solutions (e.g. SurfDrive, Sciebo)
- "ScienceMesh" by design (i.e., without app)
- MFA proposed for v1.2

SUNET

Step-up Authentication & MFA Zones

Step-up Authentication

- Adds Nextcloud-native MFA to SSO accounts

- TOTP or WebAuthn

- Can be enforced via groups (forcemfa)

- Might be made mandatory in the future for all users on all nodes

- **MFA Zone**

- Enforces MFA on file- and folder-level

- Prevents sharing with non-MFA accounts

- Works only on a local node

- Proposed OCM-extension

"The auditors tried to bypass MFA and Step-up authentication but failed. So, even though the efforts were made, our conclusion is that the quality is outstanding regarding the main target of the audit."
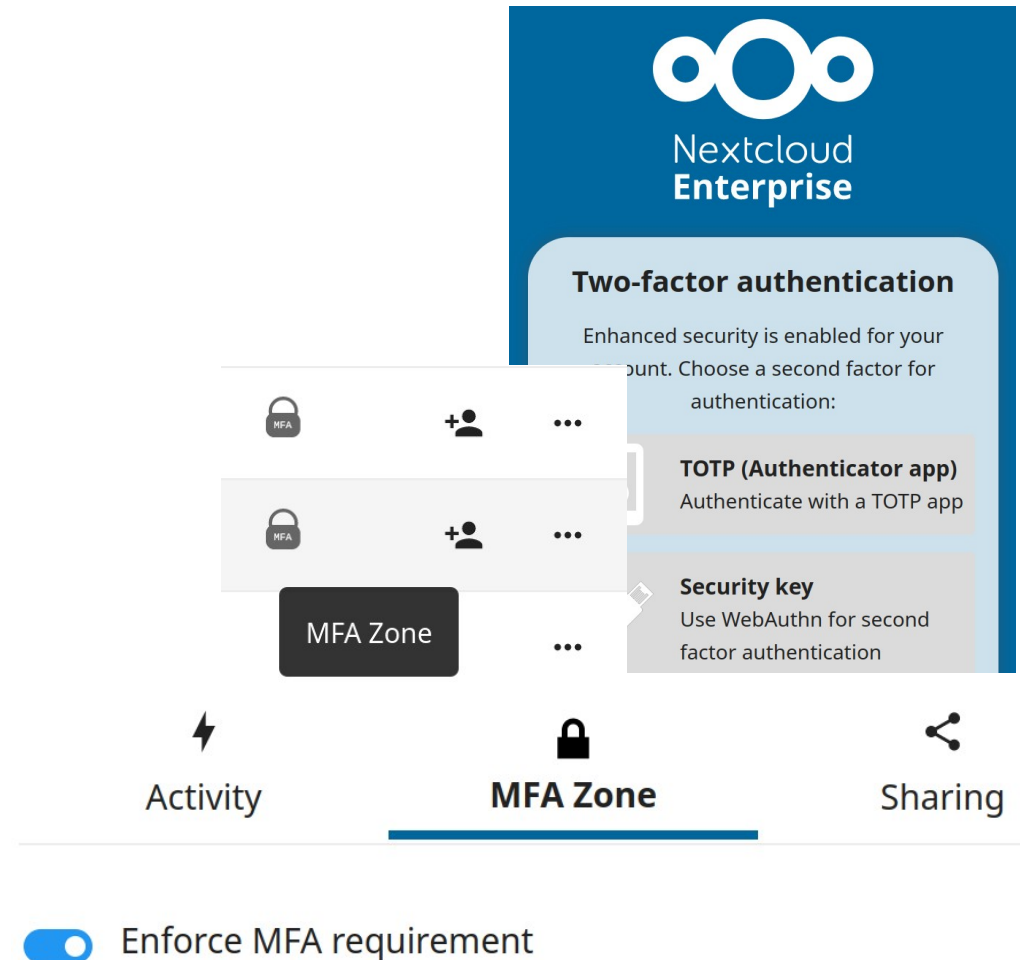
Academic Toolbox – JupyterHub

- Integrated in Sunet Drive

- Users can individually execute notebooks

- Users get 10GB "local" Jupyter Storage

- Storage is continuously synchronized between Sunet Drive and JupyterHub

- Allows people to share and collaborate on code

- Write once, execute however often you like

- Supports 40+ programming languages

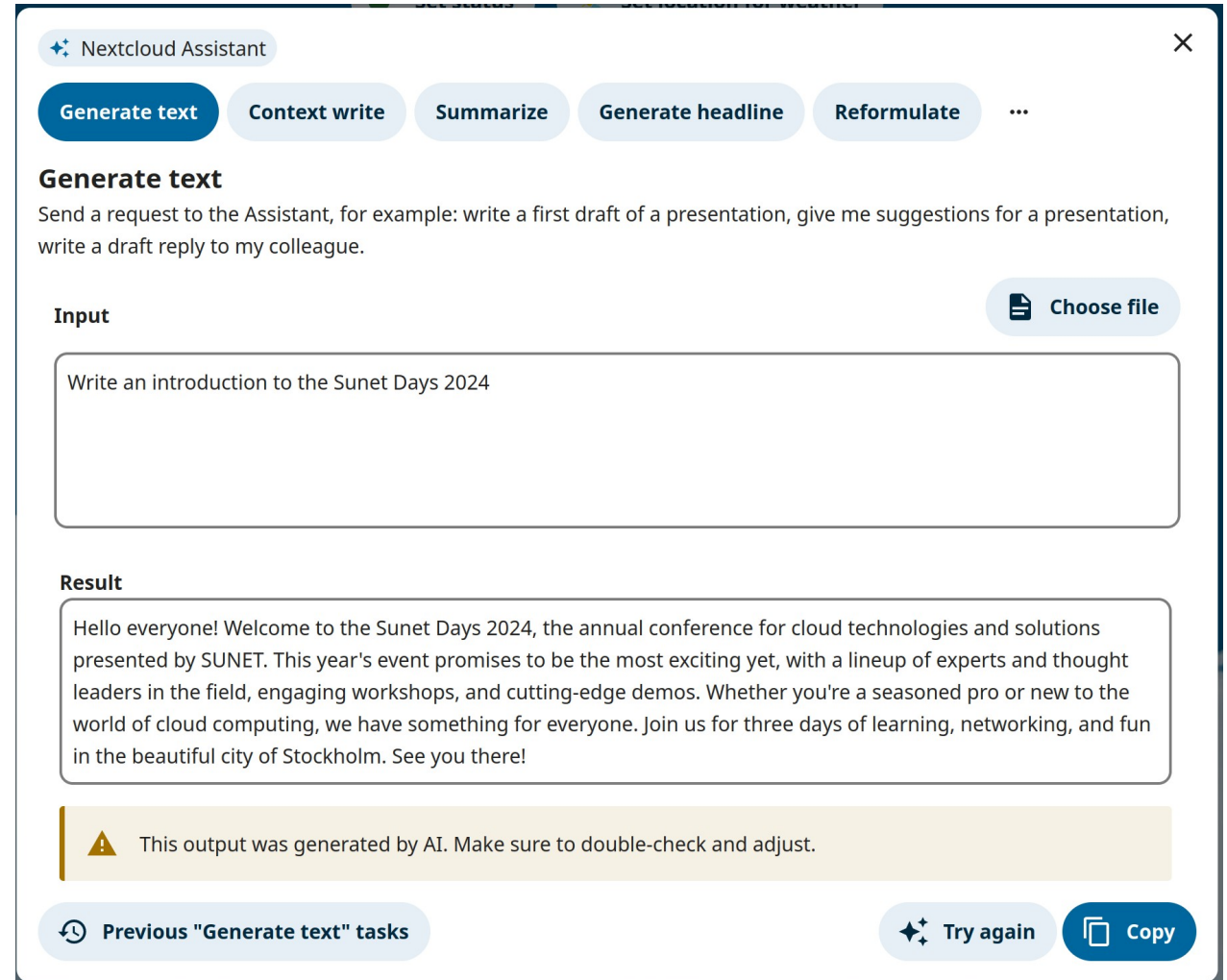- Funded by GÉANT Above-the-Net Services Incubator
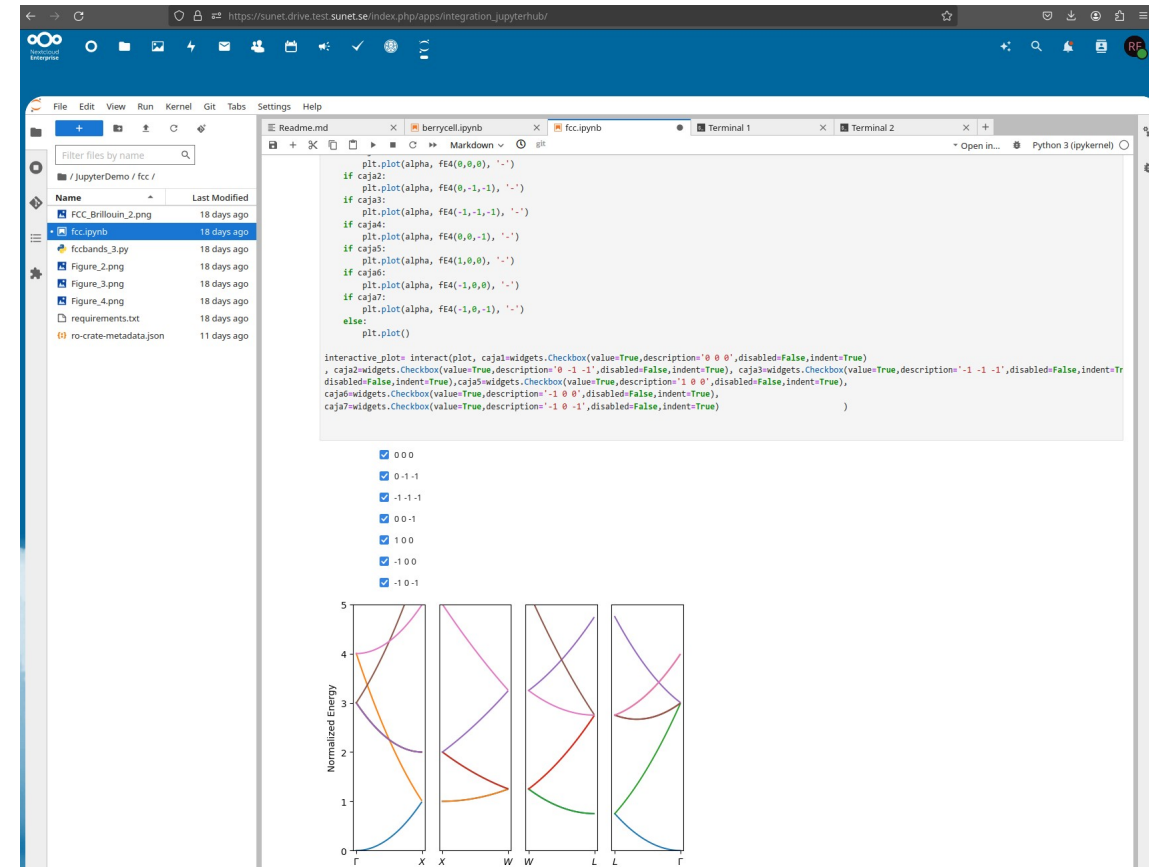
Academic Toolbox – From Data to DOI with a Berry Solar Cell

- Execute Experiment
  - Jupyter notebook with instructions to make a simple organic solar cell
  - Measure data an upload to Sunet Drive in realtime

- Data Analysis
  - Analyse and visualise data in JupyterHub

- Publish Data
  - Prepare data and metadata
  - Publish to Zenodo using **ScieboRDS** app
  - https://sandbox.zenodo.org/records/70251

- Showcased as live demo at TNC24

Berrycells.com

AI Assistants

- Introduced in Nextcloud 27.1 and improved in 28 and 29

- Has support for OpenAI and Local Large Language Models

- On-prem inference using "Ethical AI"

- LLM, Speech2Text, Image Generation and other assistants available

- Requires many resources (CPU & GPU)

- Many open questions (technical and legal)
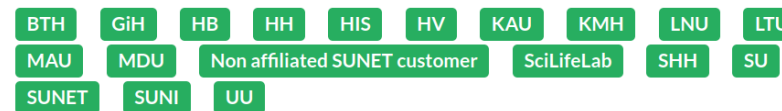
- Under evaluation

Everything else

- Open for everyone via EduGAIN

- Monitoring with Thruk: >6000 monitoring points

- Automated testing: >30k test points per day

- Infrastructure as Code

- **Open Source**

  - https://github.com/SUNET/nextcloud-custom

  - https://github.com/SUNET/nextcloud-jupyter

  - https://github.com/SUNET/nextcloud-mfazones

  - https://github.com/SUNET/drive-tests

  - https://github.com/cs3org/OCM-API

- And more: eduid.se, ScienceMesh, OCM, RDS-ng, …
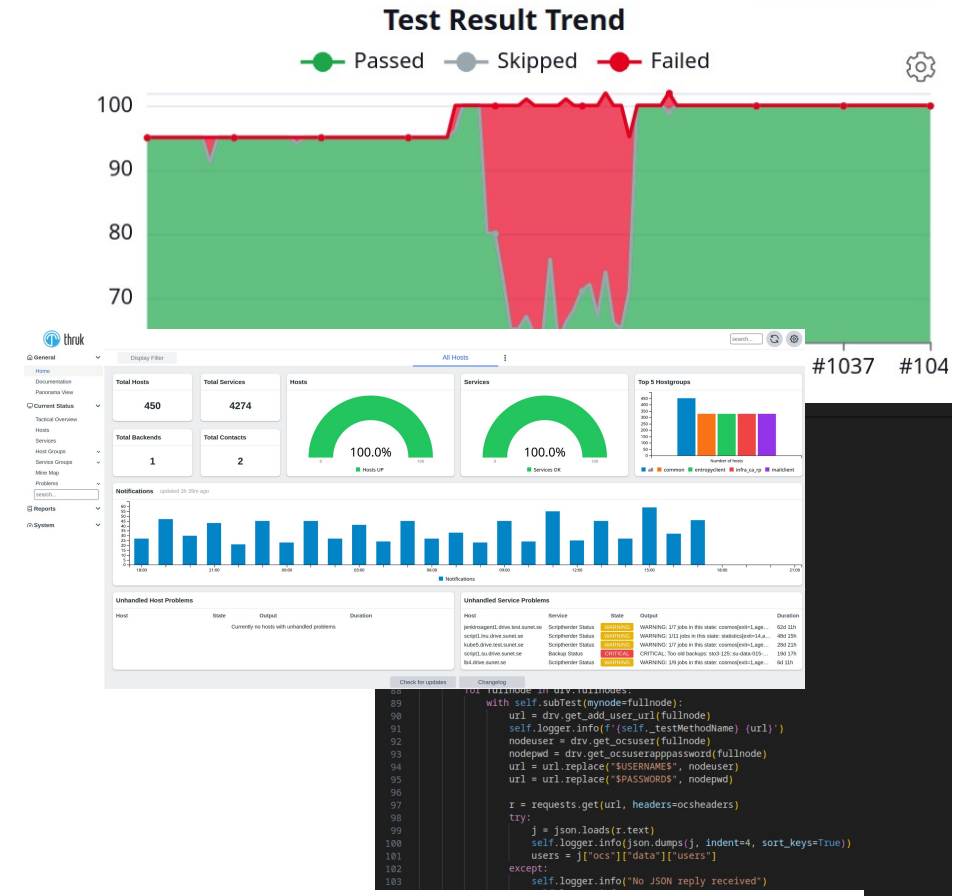
Tack! Frågor?

Anders Nilsson

anders@sunet.se

Richard Freitag

freitag@sunet.se

Micke Nordin

kano@sunet.se

Magnus Andersson

mandersson@sunet.se

Issues and Questions

drive@sunet.se

SUNET