

TSD - an eInfrastructure for sensitive data

Leon du Toit

Section Leader, Research Services

USIT



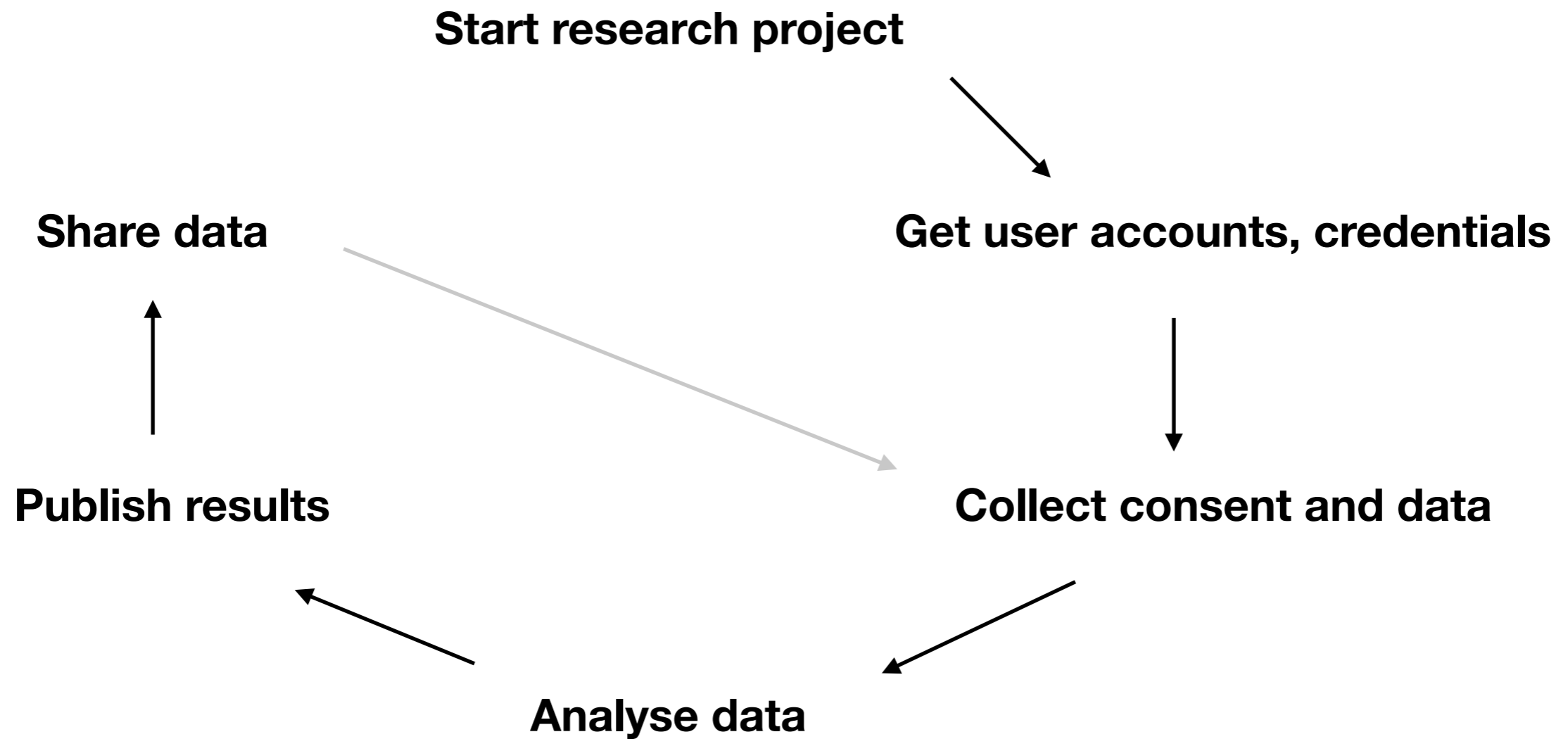
Outline

- TSD overview
- TSD and the research data life cycle
- Administrative procedures
- Data services
 - collection
 - analysis
 - storage/archival
 - sharing

TSD - classification

- Aimed at active research projects
- Enable long term archival and re-use
- Both small and large datasets
- Any discipline with sensitive data

Research life-cycle



TSD overview

- In production since 2014
- PaaS eInfrastructure for sensitive data
- Remote Desktop with Two Factor login from anywhere
- Platform for application development
- 3000+ users, 650 active projects (300 UiO)
- 3500+ user support cases per year
- National service via Sigma2: financing High Performance Computing and storage

Project creation

- REK/NSD/Datatilsynet/Personvernombud - required ethical approval
- Submission done with DIFI digital signature
- 10 working days for processing
- No limitation on who can apply as PI, or member
- Different pricing for different affiliations

TSD basic package

- 1 TB storage, with backup
- 1 Windows VM, with backup
- 1 Linux Redhat 6 VM, with backup
- Standard software
- Nettskjema (for survey/panel data collection), automated dataset construction
- HPC access if necessary (pay per use)

Project roles

- Principle investigator
 - Initial applicant
 - Responsible for project setup and rights management
- Rights
 - Admin- right to assign/remove rights
 - Export - right to export data from TSD
 - Custom - to manage custom file access

User creation

- Project admin fills in authorisation form
- Applicant fills in application
- <https://www.uio.no/english/services/it/research/sensitive-data/use-tsd/administrative-tasks/administer-users/index.html>

TSD Selfservice - Apply for use x +

localhost:5000/user-application

UiO: University of Oslo

TSD Selfservice

Selfservice portal for TSD

Language ▾

281293***** Log Out

Apply for user account

Norwegian National ID number
281293 [REDACTED]

Given names


Last name

Project code

Email address

University of Oslo username

[Submit application](#)



Contact information
E-mail: tsd-drift@usit.uio.no
Phone: 22 84 00 04

Responsible for this page.
[Services for sensitive data \(TSD\)](#)

TSD Selfservice - Pending appl x +

localhost:5000/pending-applications

UiO: University of Oslo

TSD Selfservice


Selfservice portal for TSD

Language ▾

281293**** Log Out

Pending user creation applications for "pxx"

Name	Date of birth	Email address	Action
Ola Nordmann	1982-05-17	olan@uio.no	<input checked="" type="checkbox"/> Approve <input type="checkbox"/> Reject



Contact information
E-mail: tsd-drift@usit.uio.no
Phone: 22 84 00 04

Responsible for this page.
[Services for sensitive data \(TSD\)](#)

Self service credentials management

- Login with BankID
- Change password
- Get one-time-code
- <https://selfservice.tsd.usit.no>

Data collection

- Structured data - e.g. surveys, longitudinal
- Unstructured data
- Automated data import
- Webapp/mobile development
- Dynamic digital consent
- Secure video recording
- Extensible APIs

Nettskjema

- Can use UiO or ID-porten credentials
- Must specify data formats (codebooks)
- Automatic dataset construction for use in statistical software
- <https://nettskjema.no>

Unstructured data import

- <https://data.tsd.usit.no>
 - Granular access on uploads
 - Import and Export
 - Unlimited data sizes
 - Resume
 - Available anywhere
- api.tsd.usit.no
 - Command-line client
 - Useful for automation
 - Custom app development

Automating data import

- <https://github.com/unioslo/tsd-api-client/blob/master/README.md#tsd-api-client>

Secure video recording

- Desktop app - record and stream, no intermediate storage
- Soon available on data.tsd.usit.no
- Will cover audio recording too

App development

- App developers and project managers to assist custom requests
- tsd-contact@usit.uio.no

Dynamic digital consent

- Collection and management of digitally signed consent
- Manage data access and usage in compliance with the GDPR
- For researchers: overview of who has consented to what while working with data
- For study participants: overview of what they have consented to, ability to opt out
- Everything available via APIs, for automation

Data analysis

- Windows and/or Linux VM with statistical software
- Support for custom software installation
- High Performance Computing cluster

Colossus3

- Parallel jobs
- 2176 AMD CPU cores, 17TB memory
- Apply for hours via Sigma2
- Centos and Redhat 7
- Supports Docker
- GPU nodes for machine learning
- Dragen, FPGA, special gene sequencing hardware
- <https://www.uio.no/english/services/it/research/sensitive-data/use-tsd/hpc/index.html>

Windows

- MS Office
- R, Python, Java, Perl
- SAS/STATA/SPSS
- Matlab
- NVivo Pro

Linux

- Libre Office (Open office)
- R
- Emacs, Notepad, gEdit
- Python, Perl
- All Redhat packages
- Colossus modules

Data storage/archival

- No sophistication or curation at the moment
- Plain storage on disk or tape
- 10 years or more if required
- Open for collaboration :)

Publication portal

- Share important medical findings with patients

Sharing sensitive data via metadata

- Collaborations:
 - Tryggve2, local EGA
 - EOSCHub, Secure B2SHARE
 - NSD, ambition: provide links to data

Contact

- l.c.d.toit@usit.uio.no
- tsd-drift@usit.uio.no
- tsd-contact@usit.uio.no