

Sharing sensitive data across borders

TSDs experience and plans

Leon du Toit, Section Leader, Research services (TSD), UiO Center for IT

Overview

- What is TSD?
- Cross-border data sharing (from TSDs perspective)
 - Theory
 - Practice
 - Possibilities
- Policies and protocols

TSD in brief

Special purpose eInfrastructure

- Remote login, 2FA, own IdP
- On-premises cloud, VMWare
- In production for 8 years
- 8000+ users
- 2000+ projects
- 60+ institutions
- 4000+ virtual machines
- 10 Petabytes of data
- Moving ~1 Terabyte in and out per day
- 14 full-time employees

TSD services (1)

Platform as a service

- Bring your own:
 - Commercial and data processing agreement
 - Ethical approval
 - Data
 - Software (+ licenses)
 - Hardware (we operate it for you)
 - Mobile app
 - Identity provider (for specific use cases)

TSD services (2)

Platform as a service

- Self-service
- HPC
- Survey data collection
- Dynamic digital consent
- Web, mobile app development, IoT - APIs for data collection
- Integrations:
 - Other Norwegian sensitive data infrastructures
 - Clinical systems
 - Registry data

Sharing between eInfras, how hard is it?

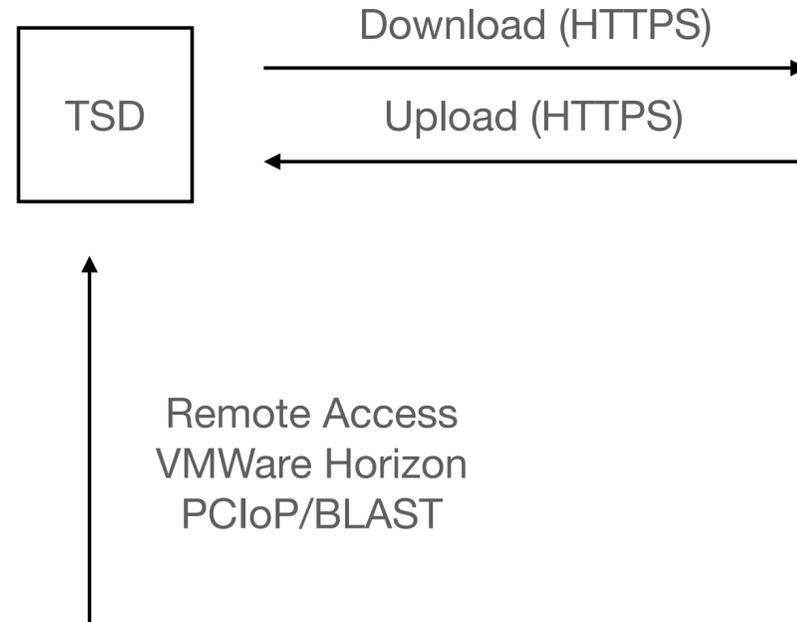
Technology, protocols, policy, geography

The world

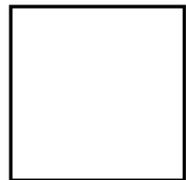
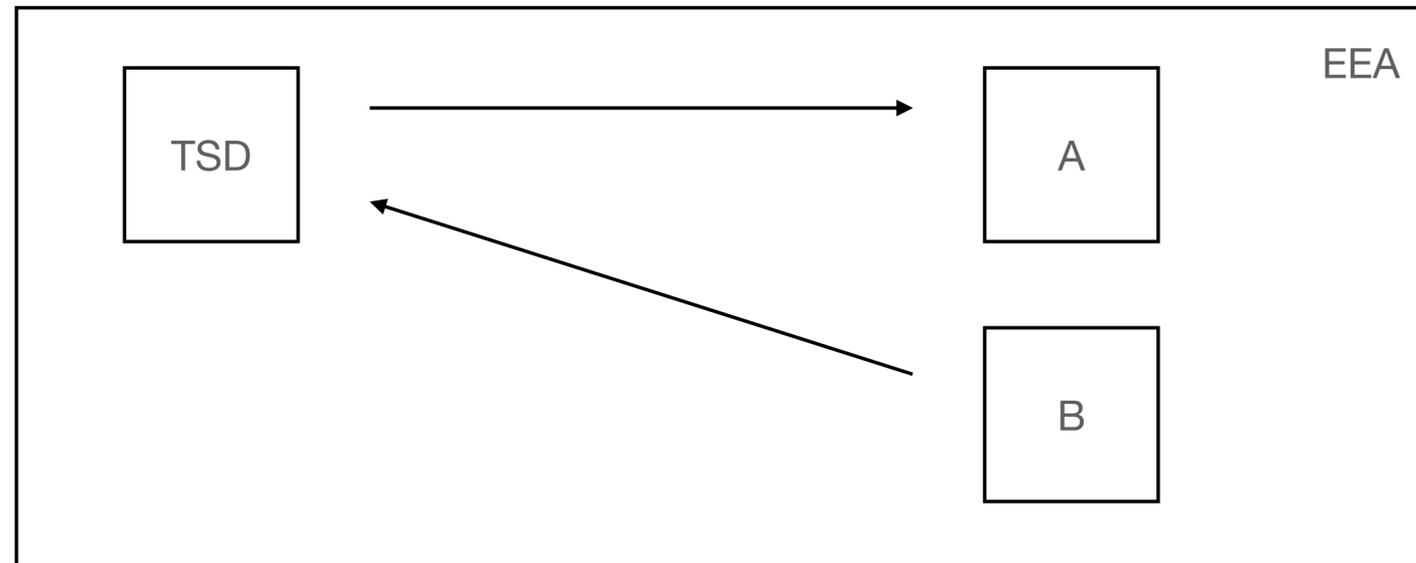
The world

TSD

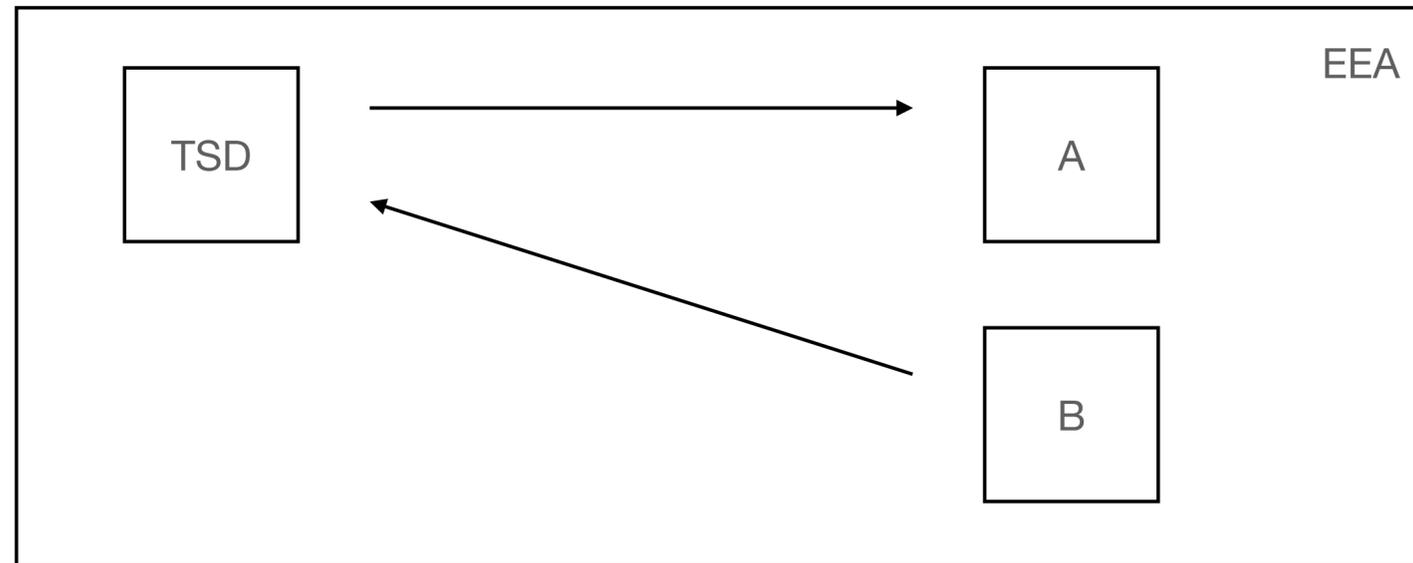
The world



The world



The world



EEA

Complicating factors

- Network/security policy
- Software installation
- Authentication
- Data volume



Solutions?

Our approach

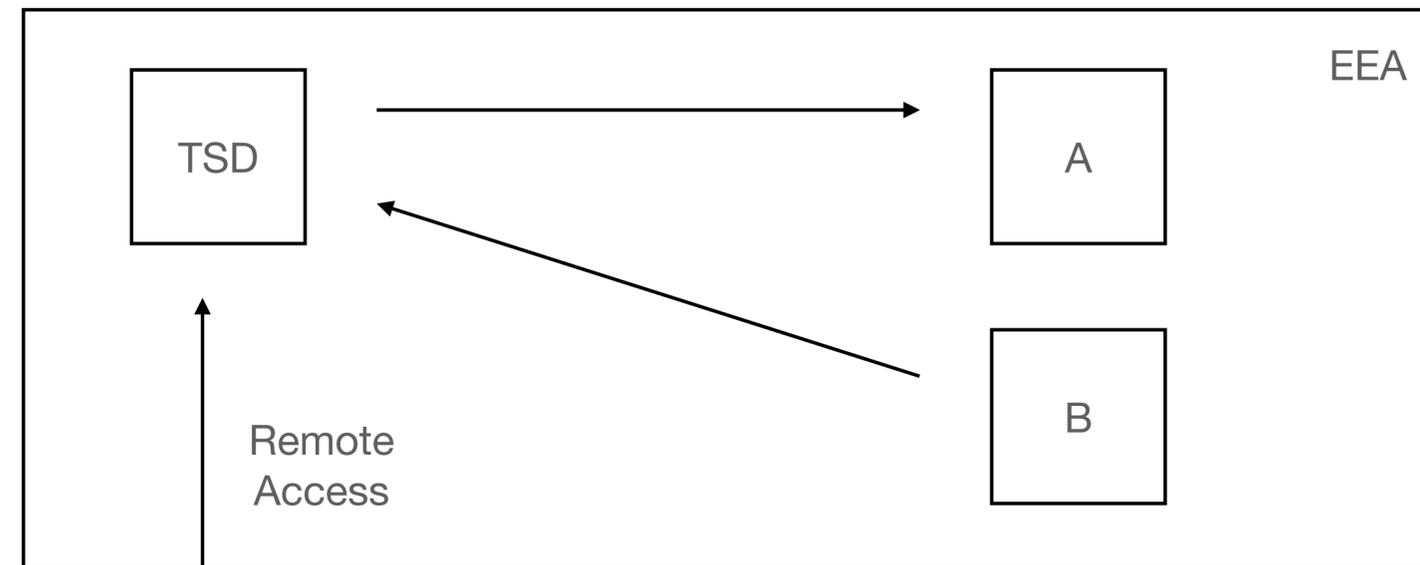
- Network/security policy
 - Make it easy to encrypt data (in addition to using HTTPS)
 - Avoid SFTP - two-way communication over ssh is problematic
- Software installation
 - A browser is enough
- Authentication
 - Import and export links - no account required
 - Associated project members (leverage eIDAS, OpenID Connect)
 - Control over our own IAM - can develop new solutions
- Data volumes
 - No limits, robust to network issues

Not enough?

Be flexible

- Use case: Norwegian Cancer Registry needed to transfer 40TB of human genomic data from UPPMAX to TSD
 - SFTP download only technical solution
- Our solution: be flexible
 - Created accounts in their research project
 - Opened the firewall temporarily
 - Download on their behalf

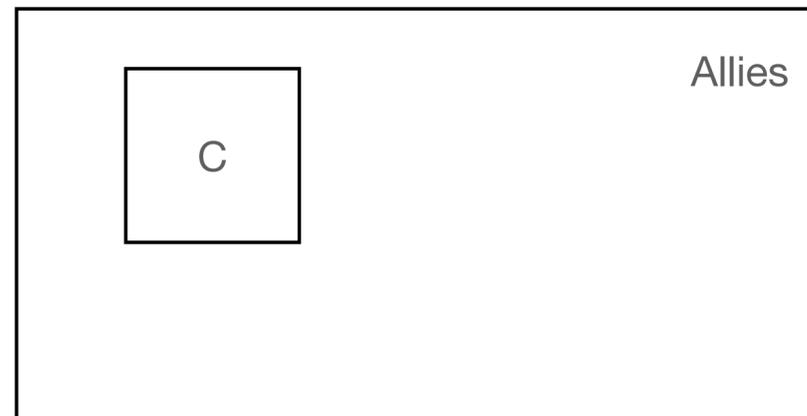
The world



EEA

Complicating factors

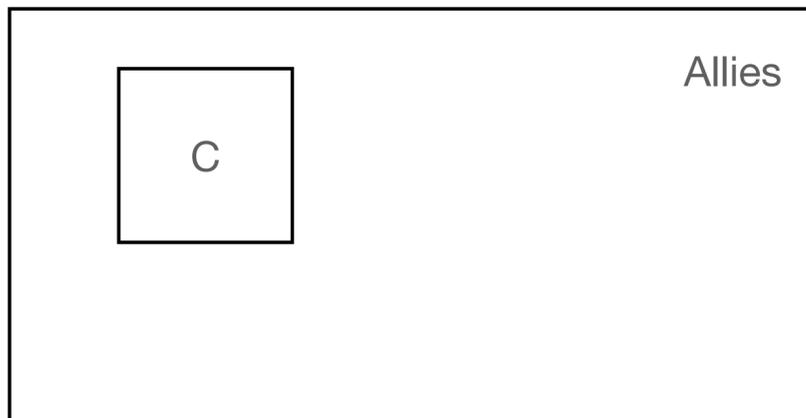
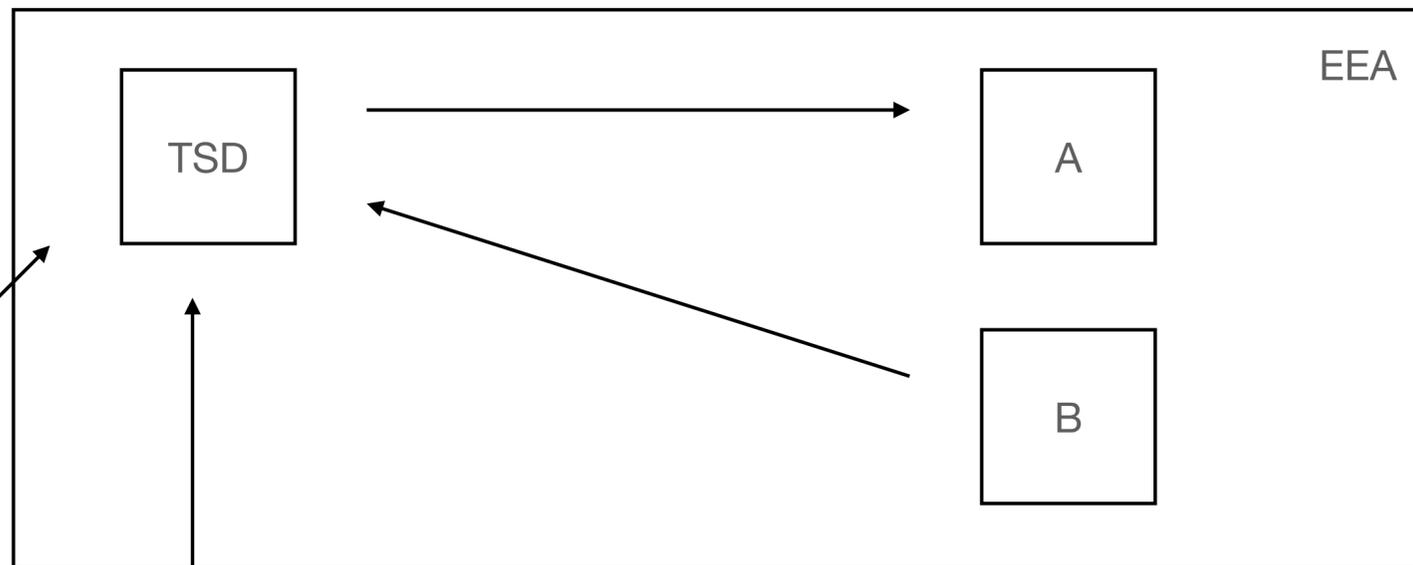
- GDPR
- Schrems ii



Allies

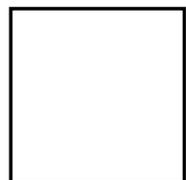
The world

The rest



Complicating factors

- Technology
- Protocols
- Policies
- Geography



Collaborations

Metadata discovery, data transport

- European Genome-phenome Archive (EGA)
- EOSC-Life
- BY-COVID
- EU-RESPONSE, Monkeypox clinical trials
- AI-Mind + EBRAINS
- UiO FAIR project

Borders

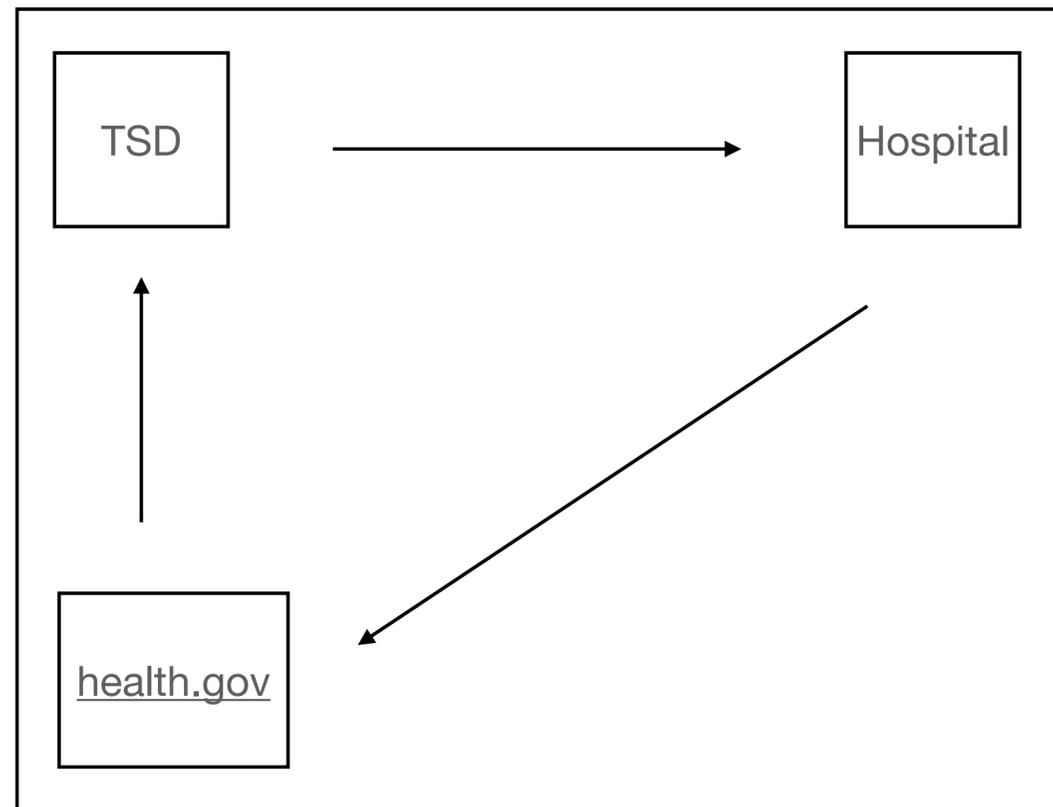
Also exist within countries

Use case

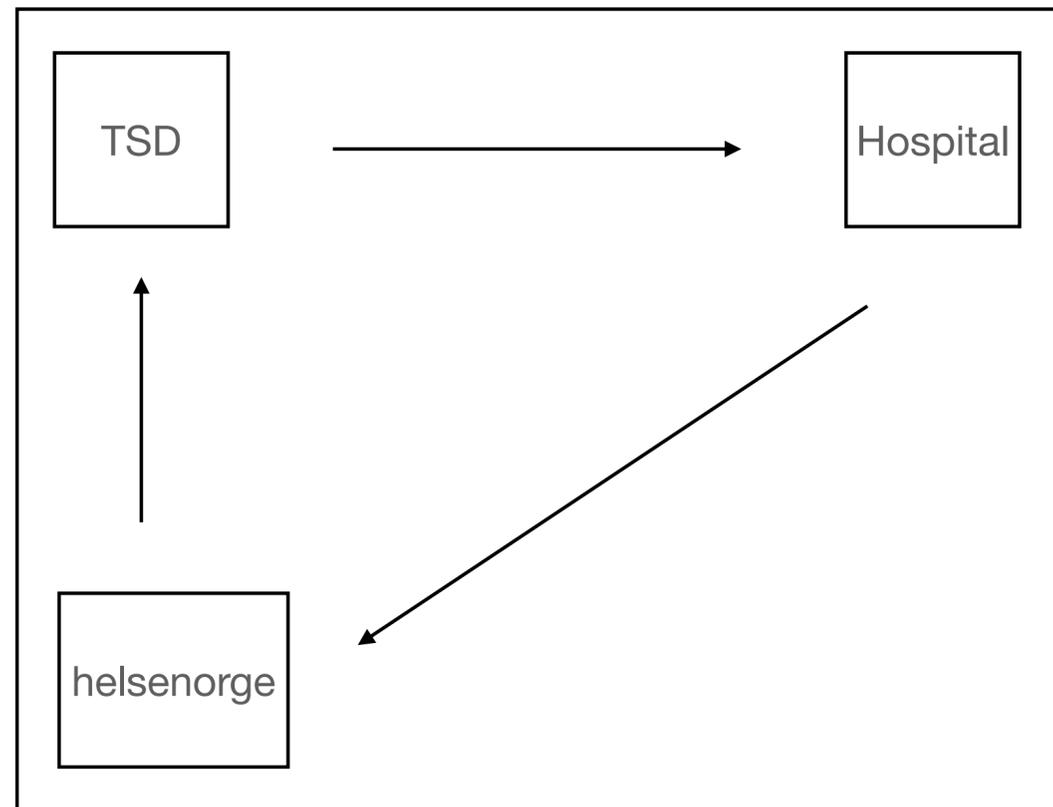
Integrating research data in clinical practice

- Clinicians at Neurology department, Oslo University Hospital want to:
 - Notify patients to complete the survey
 - Have the results in clinical systems
 - Give a copy of the data to the patient
 - Not learn to use new IT systems
- Patients want to:
 - Get help, have insight
 - Not learn to use new IT systems

The world



The world



Enabling factors:

- APIs for everything
- Using the internet for transport

Principles

- Solve specific problems
- Be flexible, do things that don't scale - within reason
- Adopt HTTPS for data transport - even if in addition to SFTP
- Define APIs for data transport - supports integrations
- Use standard protocols - like OpenID Connect
- Leverage EU regulations - e.g. eIDAS
- Employ software developers - control your own platform
- There is no silver bullet, many valid solutions

Thanks

leoncd@uio.no