HPC in the Clinic – Merge of clinical usage and research eInfrastructure

Ole Saastad, PhD
Division for Research Computing Services
University Centre for Information Technology (USIT)
University of Oslo

UNINETT

sigma2

Agenda

- Sensitive Data & GDPR
- What is TSD
- Examples with TSD and GDPR
- Clinical use of HPC

Sensitive data in the light of GDPR

~~Sensitive data~~ -> ***Special categories of personal data***

«Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.»

Except:

- **Consent, but still kept safe**
- Needed by state of employer; while still keeping data safe
- Data made manifestly public by that data subject
- **Medical treatment**, but still kept safe
- And a few more…

Ownership of the data

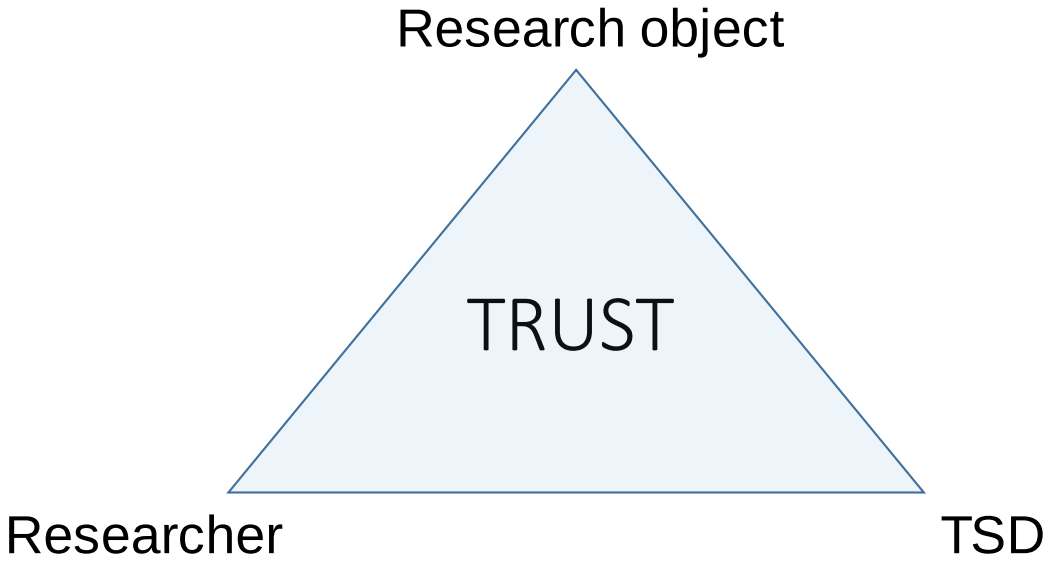
- EU commission does not really talk of ownership of this type of data;

but the right to **control the data**, and **to control whom may access and store it**

You may handle all types of data in TSD

.... if you are entitled

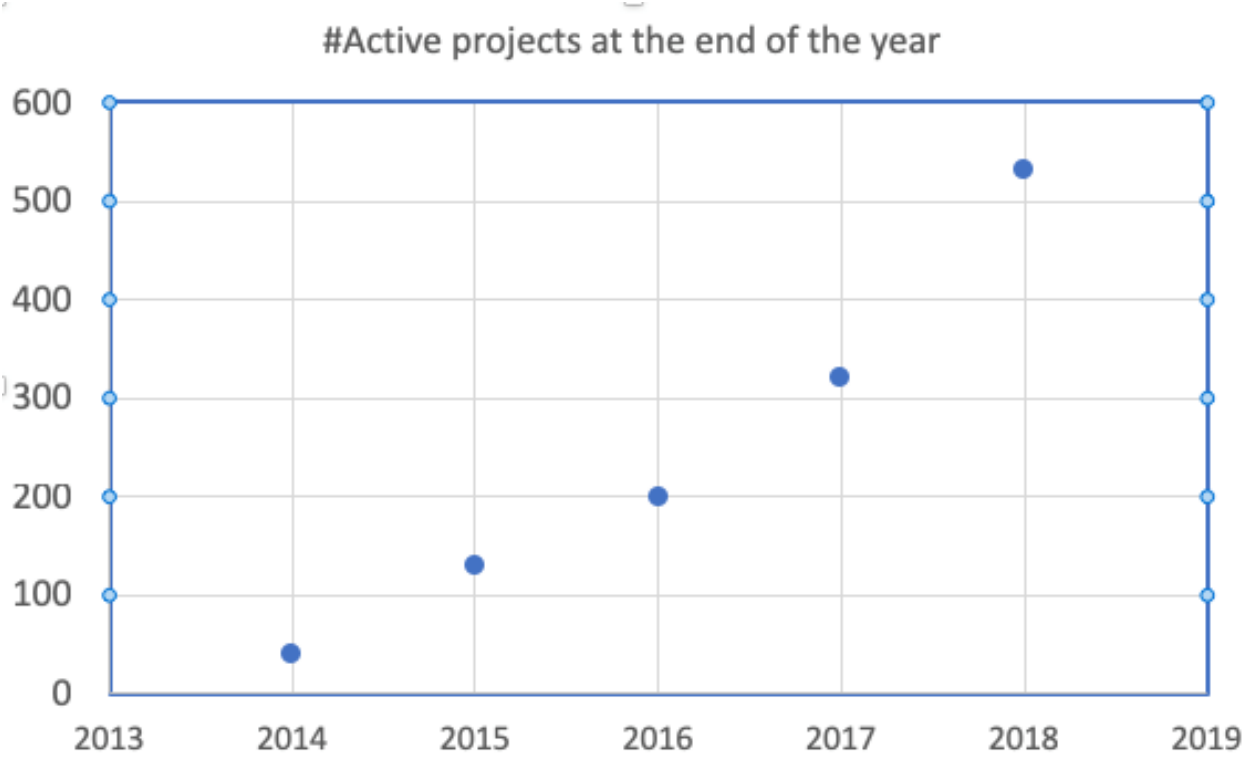
.... if you comply to the data minimization principle


Research object
TRUST
Researcher
TSD

TSD – Services for Sensitive Data

An infrastructure for high security data

- PaaS
  - remote login, with Two Factor Authentication
  - Windows and linux VMs
  - web services
- On-location data center
  - VMware
  - virtual enclaves for projects
  - layer 2 and layer 3 network separation
- HPC
  - 2000 CPU core cluster
  - GPU nodes
  - FPGA gene sequencer
- APIs for app development



> 2,5 PiB data
> 3000 users
> 2000 VMs
> 17 organisations

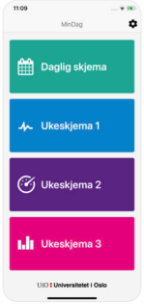
~ 10 full time positions

+ Clinical production

TSD

UiO : University of Oslo

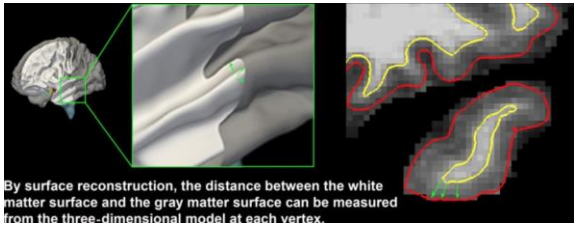
Phone apps



Self Service

Export/import API

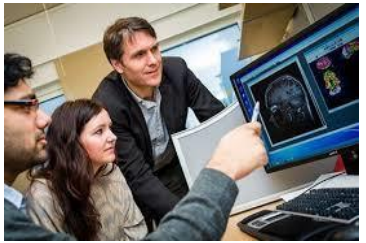
Tablet apps



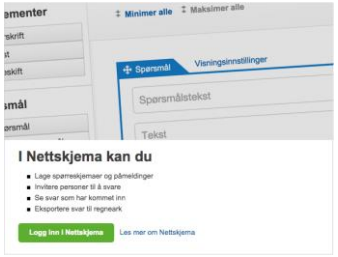
Compute, analyse and store data

TSD

Data capture



Collaborate / share

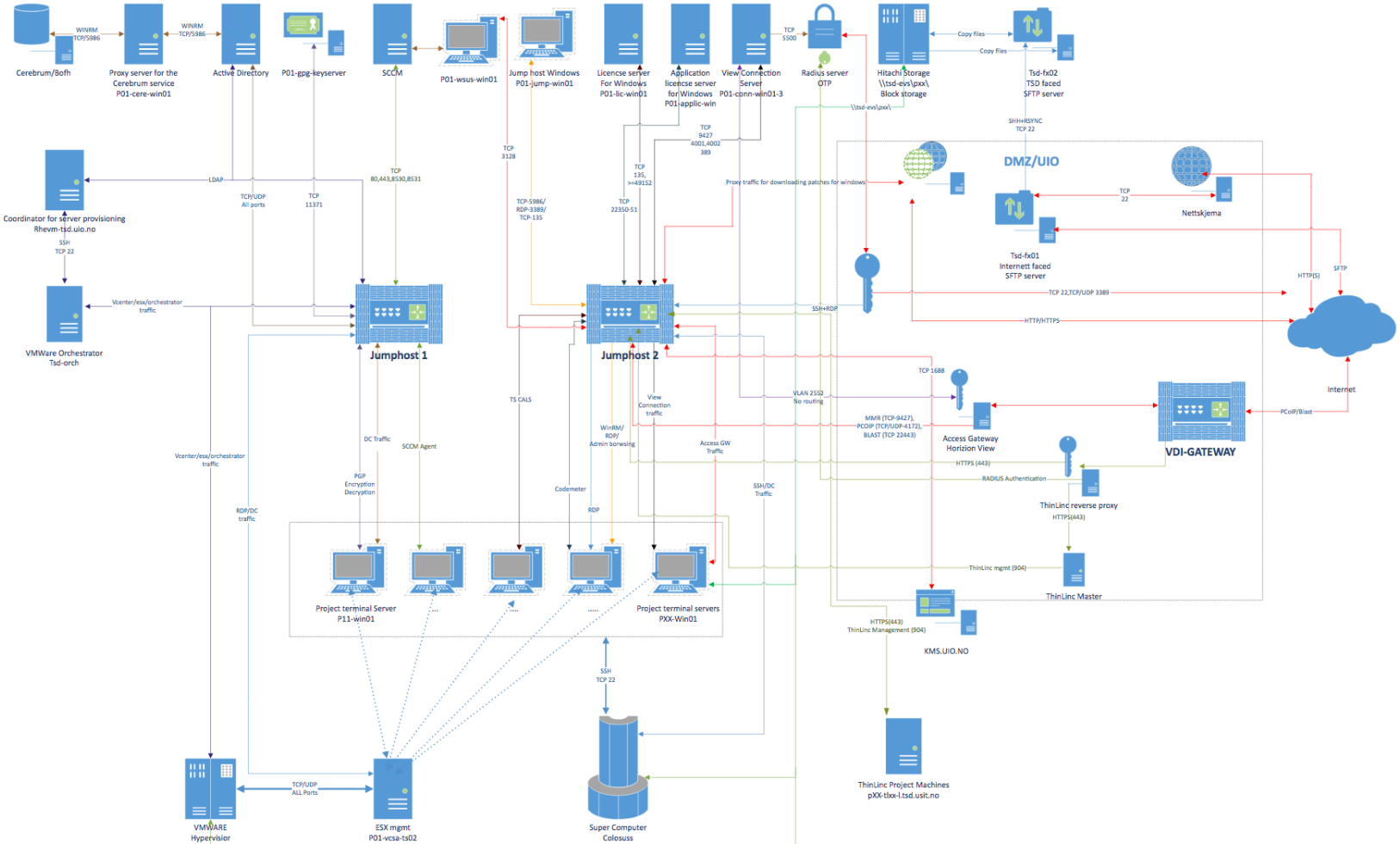


Web forms

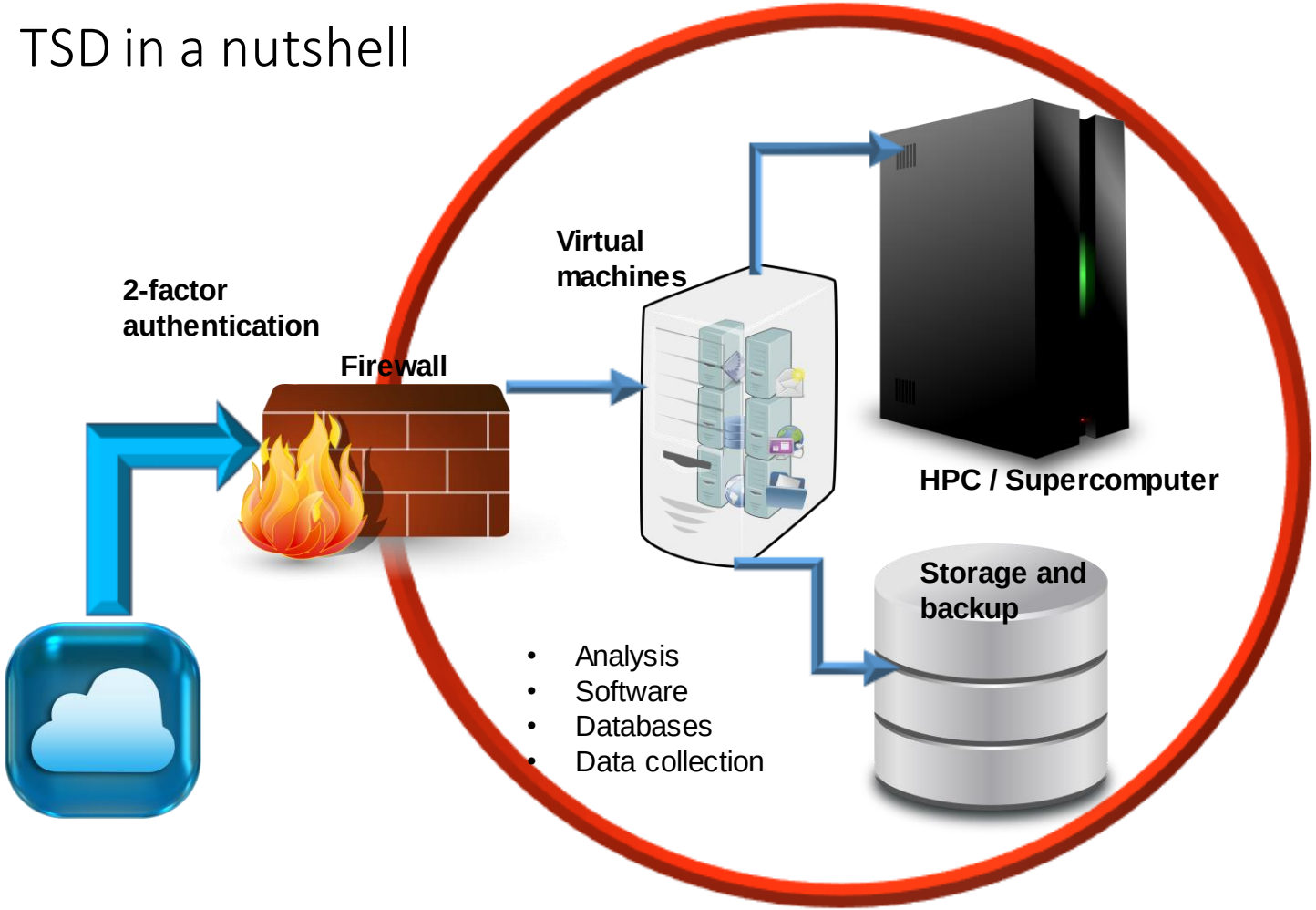
Consent Service

Analysis

Feedback services


Cerebrum/Bofh
WINRM TCP/5986
Proxy server for the Cerebrum service P01-cere-win01
WINRM TCP/5986
Active Directory
P01-gpg-keyserver
SCCM
P01-wsus-win01
Jump host Windows P01-jump-win01
License server For Windows P01-lic-win01
Application licencse server for Windows P01-applic-win
View Connection Server P01-conn-win01-3
TCP 5500
Radius server OTP
Hitachi Storage \\tsd-evs\pxx\ Block storage
Copy files
Copy files
Tsd-fx02 TSD faced SFTP server
\\tsd-evs\pxx\
SHH+RSYNC TCP 22
TCP 3128
TCP 9427 4001,4002 389
TCP 135, >=49152
TCP 22350-51
TCP-5986/ RDP-3389/ TCP-135
Proxy traffic for downloading patches for windows
DMZ/UIO
TCP 22
Nettskjema
Tsd-fx01 Internett faced SFTP server
LDAP
TCP/UDP All ports
TCP 11371
TCP 80,443,8530,8531
Coordinator for server provisioning Rhevm-tsd.uio.no
SSH TCP 22
Vcenter/esx/orchestrator traffic
VMWare Orchestrator Tsd-orch
Jumphost 1
Jumphost 2
SSH+RDP
TCP 22,TCP/UDP 3389
HTTP/HTTPS
HTTP(S)
SFTP
Internet
TCP 1688
VLAN 2552 No routing
MMR (TCP-9427), PCOIP (TCP/UDP-4172), BLAST (TCP 22443)
Access Gateway Horizion View
VDI-GATEWAY
PCoIP/Blast
HTTPS (443)
RADIUS Authentication
ThinLinc reverse proxy
HTTPS(443)
ThinLinc mgmt (904)
ThinLinc Master
HTTPS(443) ThinLinc Management (904)
KMS.UIO.NO
TS CALS
View Connection traffic
WinRM/ RDP/ Admin borwsing
Access GW Traffic
SSH/DC Traffic
DC Traffic
SCCM Agent
Vcenter/esx/orchestrator traffic
PGP Encryption Decryption
RDP/DC traffic
Codemeter
RDP
Project terminal Server P11-win01
Project terminal servers PXX-Win01
SSH TCP 22
TCP/UDP ALL Ports
VMWARE Hypervisior
ESX mgmt P01-vcsa-ts02
Super Computer Colossus
ThinLinc Project Machines pXX-tlxx-l.tsd.usit.no

TSD in a nutshell



UiO : University of Oslo

TSD

Admin services

API services

User services

Auth services

Login services

Export services

Import services

Log services

Consent service

Data discovery

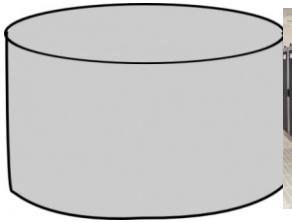
...

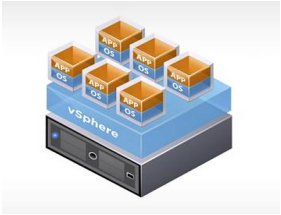
Self service

Fast dev.

Lightweight IT

Routing, firewall, integration architecture, API and authorization





Storage

HPC

VMware

Databases

Slow development

Heavyweight IT

GDPR basics

- The right to know
- The right to control and see your data
- The right to be forgotten

Human basics

- What´s in it for me?

Precision medicine in TSD – Example 1

- Oslo University Hospital – Dept of Medical Genetics
  - Rare diseases
  - Newborn screening
  - Inherited Cancer – and a few other cancer sample types

Precision medicine and IT

- Target the disease and in many cases the actual genetic and molecular footprint of the disease rather than for example treating cancer based on the tissue in which the tumor resides
- Breaks the clean distinction between research and clinic
  - One patient -> One research project
- Requires agile development platforms integrated with eInfrastructures that can handle rapid development and change

UiO : University of Oslo

Oslo universitetssykehus

Forside > Avdelinger > Klinikk for laboratoriemedisin > Avdeling for medisinsk genetikk

Avdeling for medisinsk genetikk

Avdeling for medisinsk genetikk er landets største medisinsk genetiske avdeling og arbeider med utredning av arvelige sykdommer og forskning på arvelige årsaker til sykdom.

Avdeling for medisinsk genetikk

Avdeling for medisinsk genetikk er landets største medisinsk genetiske avdeling og arbeider med utredning av arvelige sykdommer og forskning på arvelige årsaker til sykdom.

Hovedvirksomheter ved avdelingen er klinisk genetisk utredning, genetisk veiledning inkludert veiledning i forbindelse med fosterdiagnostikk, genetisk laboratoriediagnostikk og genetisk forskning.

Avdelingen ledes av professor Dr.med. Dag Undlien, og er organisert i seksjoner. Laboratorievirksomheten er samlokalisert ved Ullevål sykehus og den kliniske virksomheten ved Rikshospitalet

Avdelingens drift utføres i samsvar med Lov om medisinsk bruk av bioteknologi.

Clinical use
Clinical use
Clinical use

AMG - OUS

TSD

Issues at Oslo Univ. Hospital

- No HPC
- No agile development
- No storage suited for these data types
- No support on systems suited for DNA-sequencing

Clinic versus research – admin challenges

- Special data handler agreements
- SLA wanted at 99.999% uptime
- If TSD is down 4 hours demands extra crisis measurements at the hospital
- TSD personell not contracted for work off regular hours
- No financial room for test and development platform

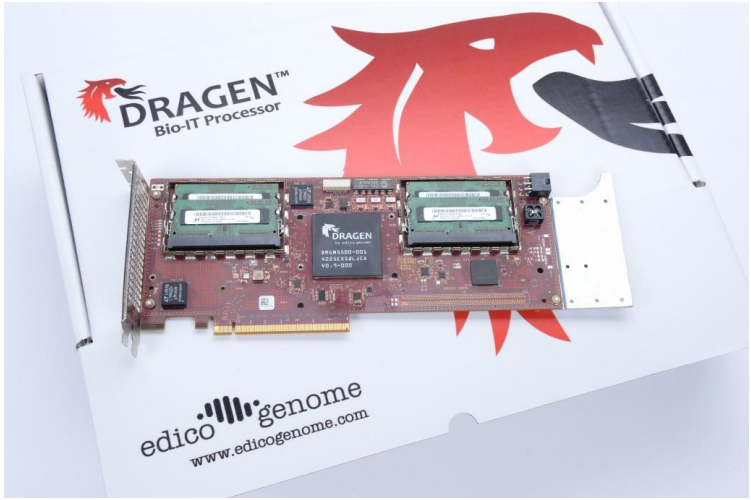
TSD delivers to Dept of Medical Genetics:

- Storage for approx 0,5 PiB of exome sequencing data
- HPC capacity for 6000+ exomes per year (moving to full genomes)
- An internal web application where medical geneticists do the mutation assessement and diagnosis (50-80 active users every day)
- Long time storage of all data
- Excessive logging and auditing
- Versions of all software used in the clinic
- Diagnostic reports are copied back to the EPJ

Dragen FPGA accelerator for special samples

When in a hurry use accelerators, 22 minutes on one full genome.

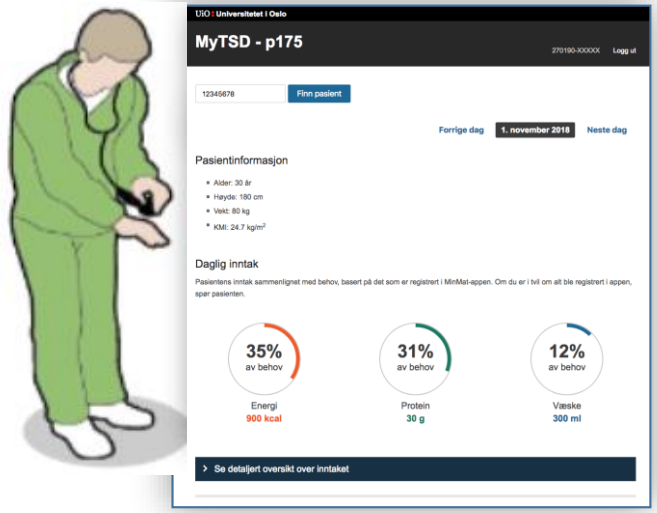
HPC walltime approx 22 hrs with todays pipeline.



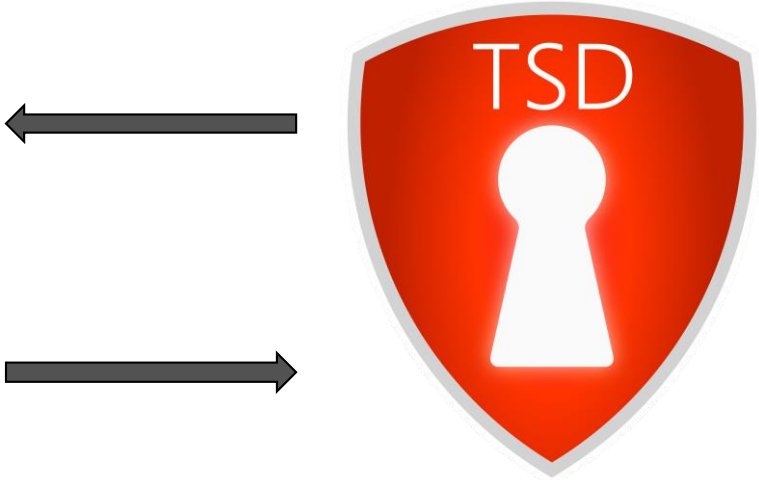
Technical challenges

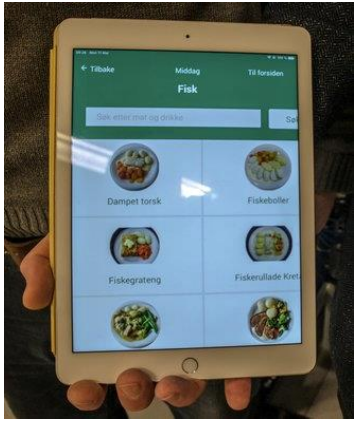
- Millions of small files – high IO load and metadata load
- Nasty neighbour problem – HPC is a shared resource QoS
- Windows VMs must access results that resides on BeeGFS
- Software updates, patching etc inside a «closed environment»
- Secure data transport
  - From DNA sequencers to TSD
  - From TSD back to Electronic Patient Journal (EPJ)

Precision medicine example 2



MinMat project at Rikshospitalet

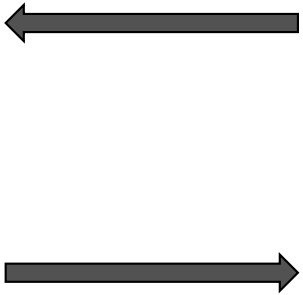




Precision medicine example 3

LetsGo-project at «Sørlandet» hospital, 2019

App with AI



Precision medicine example 4

Kidney transplant monitoring

UiO : Realfag og teknologi

Titan.uio.no
...ny kunnskap hver dag!

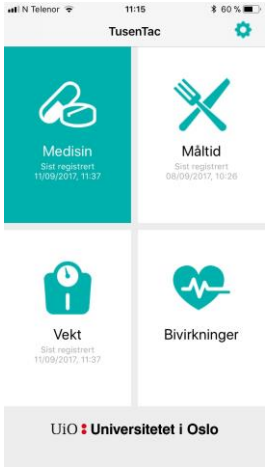
– Det er veldig individuelt hvor mye immundempende medisin ulike pasienter trenger. Derfor er det veldig viktig å finne den riktige dosen til hver pasient, forteller Anders Åsberg ved Farmasøytisk institutt. Bruk bildet.

App for transplanterte

permalenke

Folk som har fått transplantert nye organer må ta en nøyaktig dose immundempende medisiner to ganger daglig resten av livet. Farmasøytisk institutt og USIT skal nå utvikle en mobil-app som kan hjelpe pasientene med å huske på de livsviktige medisinene.

av Gunhild M. Haugnes – 11. november, 2015



https://titan.uio.no/node/983

TSD

Strategic and development roadmap

- strategic
  - consolidate as national service via public institutions
  - enhance interoperbility with other European sensitive data eInfrastructures
  - project participation
  - continue rollout of application development platform
- GDPR related development
  - full feature implementation of the GDPR for survey data
  - enforce consent on collection and usage
  - right to insight
  - right to access
  - right to be forgotten

National, Nordic, European collaboration

- Sigma2
  - funded by Dept. of Education and Research
  - project owner: involves all Norwegian Universities
  - partner in Nordic and European projects
- Tryggve2
  - Nordforsk/NeIC funded project
  - Norway, Sweden, Denmark, Finland
  - sensitive data use cases in the Nordics
- EOSCHub
  - TSD in the portfolio of services
  - developing metadata service for sensitive data
  - containers for reproducible science
- PRACE
  - UiO leads the containers Work Package

More info

- TSD White Paper, and other enquiries: tsd-contact@usit.uio.no
- https://www.uio.no/english/services/it/research/sensitive-data

Thank you for the attention

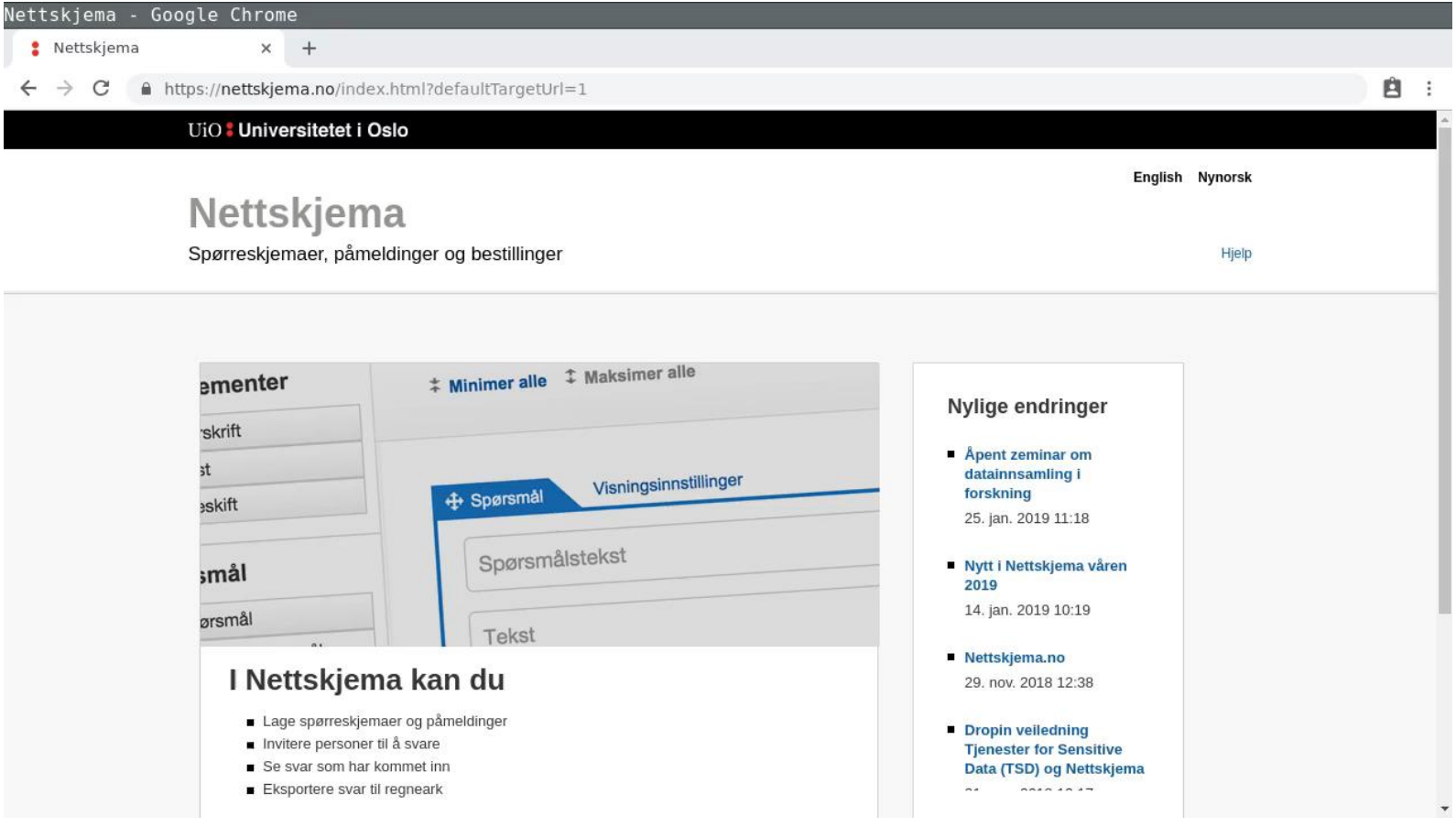
- Anyone may use our services!

https://uio.no/tsd
https://www.eosc-hub.eu/catalogue/Services%20for%20sensitive%20data

tsd-contact@usit.uio.no

UiO : University of Oslo

Digital dynamic consent system



TSD