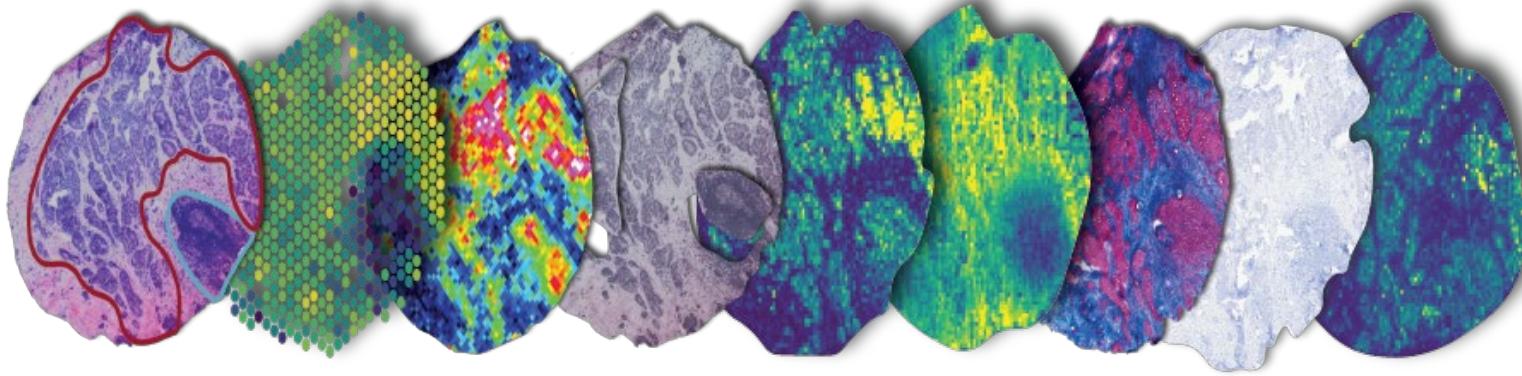


All about secure data archiving Part 3: FEGA Norway - For human genomic & transcriptomic data



Federated
European
Genome-phenome
Archive
Norway





Depositing the ProstOmics multi-omics datasets with FEGA Norway

Maria K. Andersen
Researcher, NTNU

Why we needed to archive..

- Required by funder (ERC grant no. 758306)
- Challenging for 'sensitive' health/personal data
 - Very few repositories with restricted access + approved security levels
- But final push: journal required it..



European Research Council
Established by the European Commission



COMMUNICATIONS
BIOLOGY

SPRINGER
NATURE

The days one could get away with this statement is over..

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, M.G., upon reasonable request.

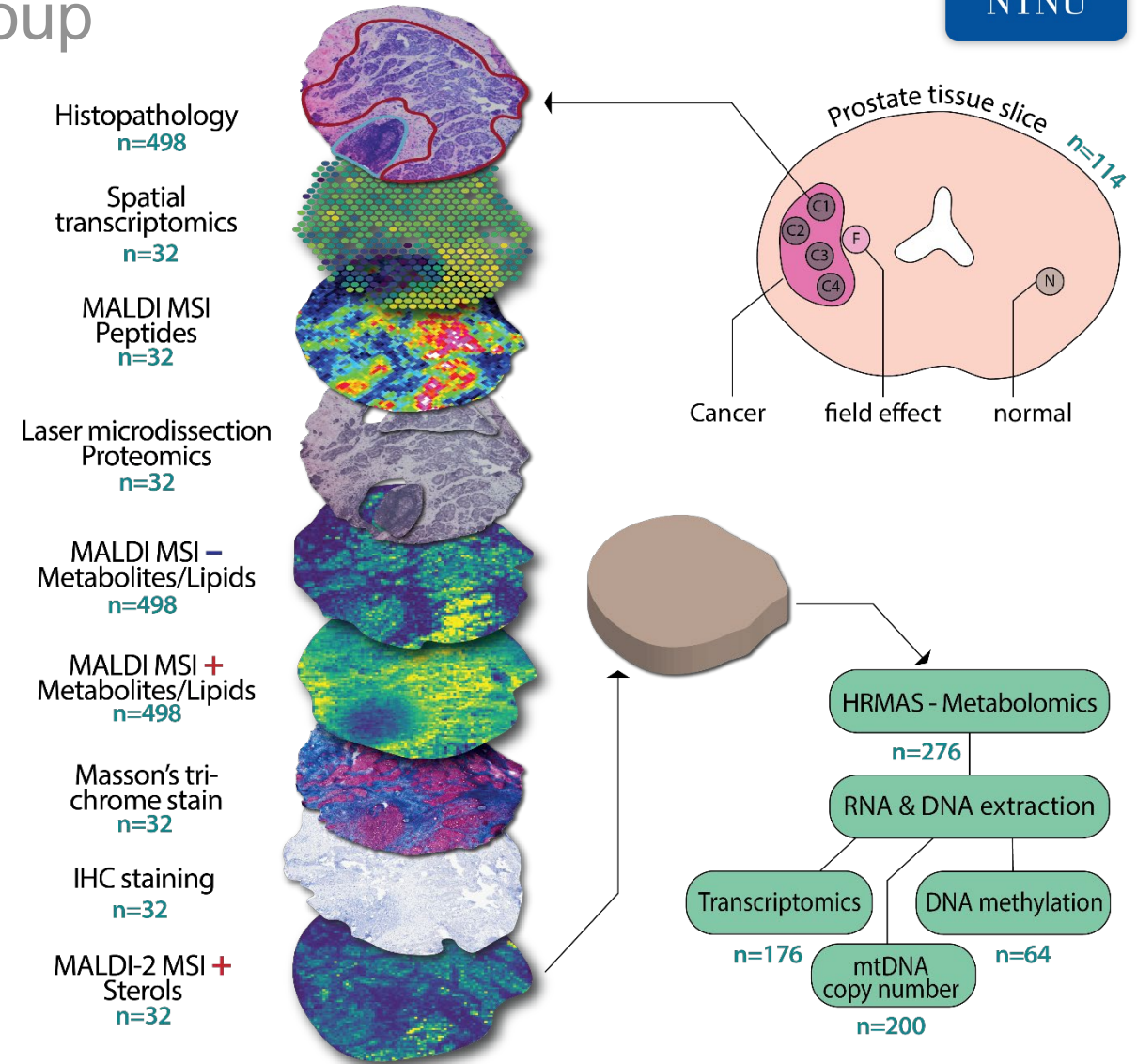
Spatial **multiomics** for prostate cancer profiling

- The core mission of ProstOmics group

Several
analytical
measurements
on each
sample

~ 500 fresh frozen prostate
samples from 114 patients

- Metabolomics
- Lipidomics
- Proteomics
- Transcriptomics
- Epigenomics
- Histopathology



Spatial **multiomics** for prostate cancer profiling

- The core mission of ProstOmics group

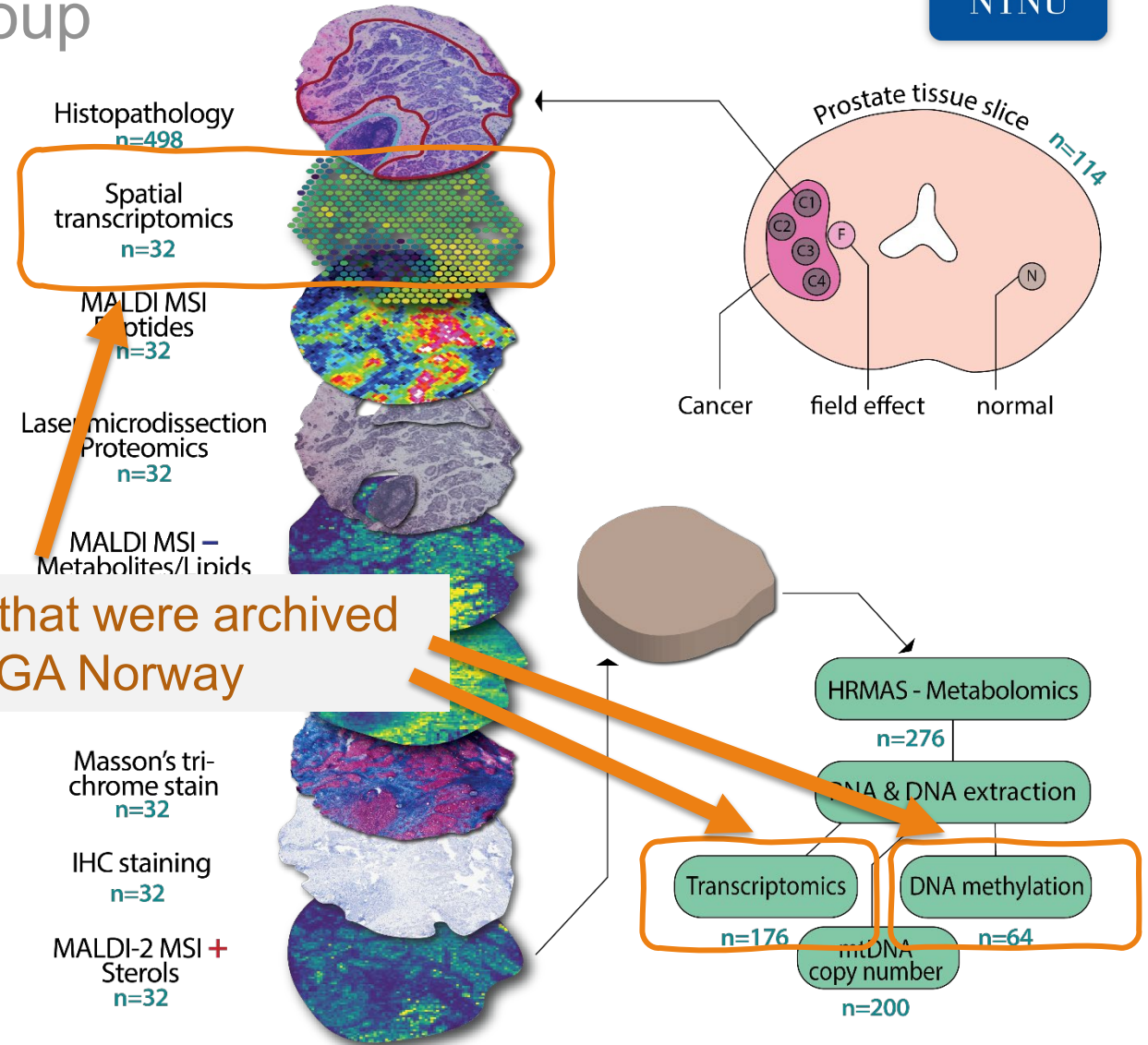
Several analytical measurements on each sample

~ 500 fresh frozen prostate samples from 114 patients

- Metabolomics
- Lipidomics
- Proteomics
- Transcriptomics
- Epigenomics
- Histopathology

Also, tried to get FEGA Norway to accept our other omics data.. But not accepting

Data that were archived in FEGA Norway



The steps

1. Paper work



2. Getting into portals



3. Prepping Metadata & data transfer



4. Processing data requests



Updating our DPIA

DPA w. NTNU & UiO
Creating a DAC
Writing a DAP

Physically: Getting EGA-profiles
Mentally: Learning how it works

There are good guidelines,
but are a bit hidden
Got a lot of help from
personell at Elixir



Kjell Petersen
Elixir, UiB

Metadata

Study description
Dataset descriptions x4
Experiments
Variable descriptions

Submitting data

Requires computational skills
Command line + scripting

We will get back to this
never-ending cycle



Maximilian Wess
NTNU

Computer
guy

DPIA= data privacy impact assessment
DPA = data process agreement
DAC = data access committee
DAP = data access policy

Studies → EGAS50000000413

One study

ProstOmics: spatial and bulk multi-omics of prostate cancer

In this ERC-funded ProstOmics project we have used spatial and bulk multi-omics on fresh frozen prostate cancer samples to investigate cancer biology and find for biomarkers to improve patient treatment. All tissue material were collected from prostate cancer patients undergoing radical prostatectomy who had not received any prior cancer-specific treatment. Of the 498 prostate tissue samples (114 patients) included in our project, 176 samples (N=37 patients) have been analyzed with bulk transcriptomics (RNA-seq). Some of these 176 samples were also analyzed with spatial transcriptomics (n=32, N=8, Visium 10x RNA-seq) and DNA methylomics (n=96, N=24, array). All datasets include metadata for histopathological evaluation. Patient metadata include information of age at surgery, time (months) until reported relapse, pre-surgery PSA and post-surgery T-stage.

Type: Transcriptome Sequencing

Archiver: Federated European Genome-Phenome Archive (FEGA Norway)

4 Datasets

5 Publications

Click on a Dataset ID in the table below to learn more, and to find out who to contact about access to these data

Dataset ID	Description	Technology	Samples
EGAD50000000603	Spatial transcriptomics data (ST) from 32 human prostate tissue samples ori...	NextSeq 500	32
EGAD50000000604	Bulk transcriptomics data from 176 prostate tissue samples (37 patients, 27 ...	NextSeq 500	176
EGAD50000000605	Bulk methylation array data from 64 prostate tissue samples from 16 patient...	Infinium MethylationEPIC BeadChip	64
EGAD50000000606	Bulk methylation array data from 32 prostate tissue samples from 8 patients ...	Infinium MethylationEPIC v2.0 BeadChip	32

4 datasets

Data access Committee



Tone Bathen
(Professor, ISB)

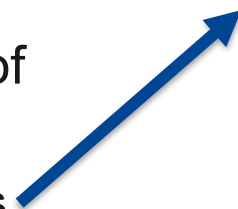


ERC PI: May-Britt Tessem
Prof. ISB, NTNU



Maria K. Andersen
(Researcher)

- >40 Data requests (~20 people)
- Made templates based on location of requesters:
 - EU and GDPR-compliant countries
 - United states
 - Other countries
- None shared through FEGA Norway yet..
 - Bottleneck: paperwork and legal assessment



Dear [Researcher's Name],

Thank you for your expressed interest in our [omics] data through the European Genome-Phenome Archive (EGA) portal.

Please be aware that the process to gain access to this sensitive data will be extensive due to the strict regulations on such data by both GDPR, the Norwegian law and the specific patient consent used for the tissue donors in our cohort.

1. We will need to assure that your scientific purpose for using these data and data handling is compliant with GDPR and the specific ethical approval.
2. We will have to apply to our regional ethical committee for approval to share the data with you
3. We will likely be required to update our Data Protection Impact Assessment (DPIA), which needs to be approved by our hospital
4. There then needs to be a data transfer agreement and collaboration agreement between NTNU and your institution, detailing conditions such as who will have access to the data, what the data will be used for and co-authorship in the event of publishing results generated from these data.
5. You will likely be required to work on this data remotely from HUNT Cloud (<https://about.hdc.ntnu.no/>), an ISO-certified cloud system for sensitive data. You will have to set up further agreements with HUNT Cloud and cover the cost to have a working environment there.

This whole process will take several months and substantial work resources from both parties. Please let us know if the required work to get access to this data is compatible with your research plan ahead.

Kind regards,
Maria K. Andersen, PhD
On behalf of the data access committee

Sharing 'sensitive' research data

Sharing with a requester in a GDPR compliant country (whether through FEGA Norway or not)

1. Is sharing data in line with: a) ethical approval and b) info given to participants?

2. How to give access?
Do they have a secure place to work on?



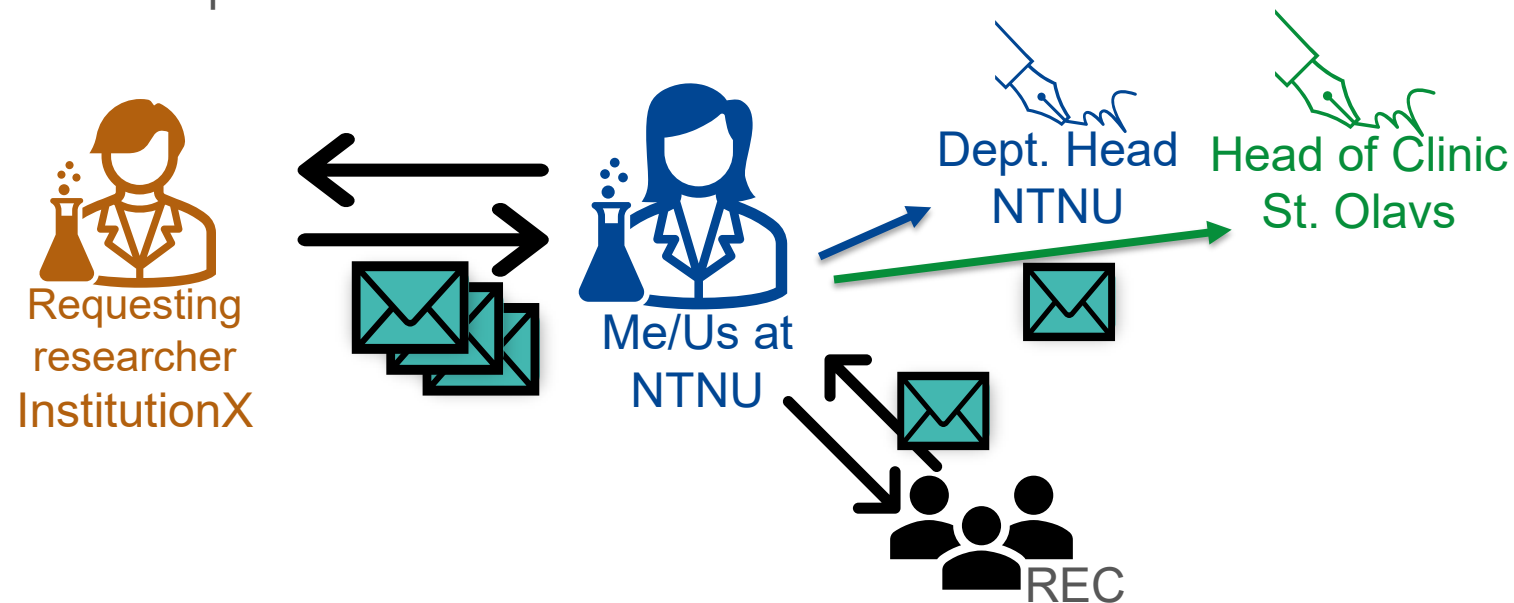
Sharing 'sensitive' research data

Sharing with a requester in a GDPR compliant country (whether through FEGA Norway or not)

1. Is sharing data in line with: a) ethical approval and b) info given to participants?

2. How to give access? Do they have a secure place to work on?

3. Apply REC +
4. Update DPIA



Sharing 'sensitive' research data

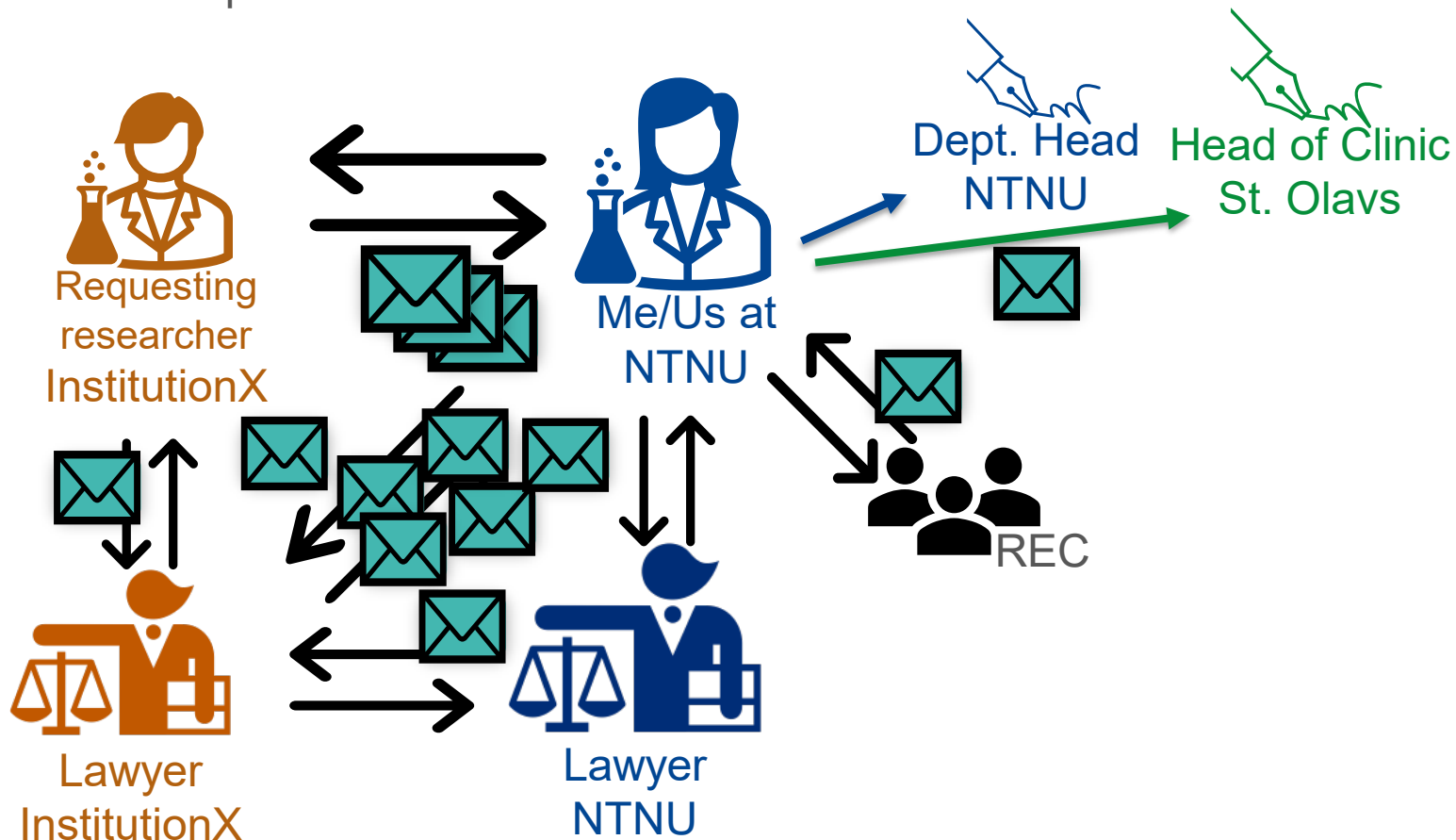
Sharing with a requester in a GDPR compliant country (whether through FEGA Norway or not)

1. Is sharing data in line with: a) ethical approval and b) info given to participants?

2. How to give access? Do they have a secure place to work on?

3. Apply REC +
4. Update DPIA

5. Data transfer agreement + collaboration agreement



Final thoughts:

The good

- Great with an infrastructure like FEGA Norway; a repository with restricted access
- More visibility to our research
- Subsequent publications on the same data made easier
- Got «forced» to make the data FAIR
 - Enabling for reuse with metadata and readme-instructions.
 - Handy for new people joining our group

And the ‘bad’

- Hoping for more flexibility in FEGA Norway
 - For other omics data and clinical data.
- Bottle-neck: All the legal privacy regulations
 - (not the fault of FEGA Norway though)
 - The processing and legal requirements for sharing data with other researchers is still unreasonable a lot of work

**Thank you for
your attention!**

