

Infrastructures for storing, sharing and archiving data

Where do we store and preserve qualitative data today?

Results from an UiO survey

Chief Engineer Torgeir Christiansen, Teaching Learning Video lab (TLVlab), Department of Teacher Education and School Research, UiO

Dr Anders Lunde, Oslo Special Education and Learning Lab, Department of Special Needs Education, UiO

Dr Agata Bochynska, Open Research, University of Oslo Library

What infrastructure solutions can we provide for qualitative research at UiO?

Dr Anne Schad Bergsaker, IT-department, UiO



QualiFAIR frokostseminar 14.05.2024

QualiFAIR Infrastructure Working Group

Torgeir Christiansen, Teaching Learning Video lab (TLVlab),
Department of Teacher Education and School Research, UiO

Anders Lunde, Oslo Special Education and Learning Lab,
Department of Special Needs Education, UiO

Kristin Hagen, Text laboratory and Humit, UiO

Agata Bochynska, Open Research, University of Oslo Library

Anne Schad Bergsaker, IT-department, UiO

Survey about solutions for storing and sharing qualitative data

0 %



Mandatory fields are marked with an asterisk *

Would you like to influence what storage solutions are offered for qualitative data at UiO? Do you want to help us make it easier to share or reuse qualitative data?

We would greatly appreciate it if you could use **5-10 minutes** of your time to fill out this form. Here, we gather information about solutions and challenges in storing and reusing qualitative data as well as needs for further development of solutions for data at UiO.

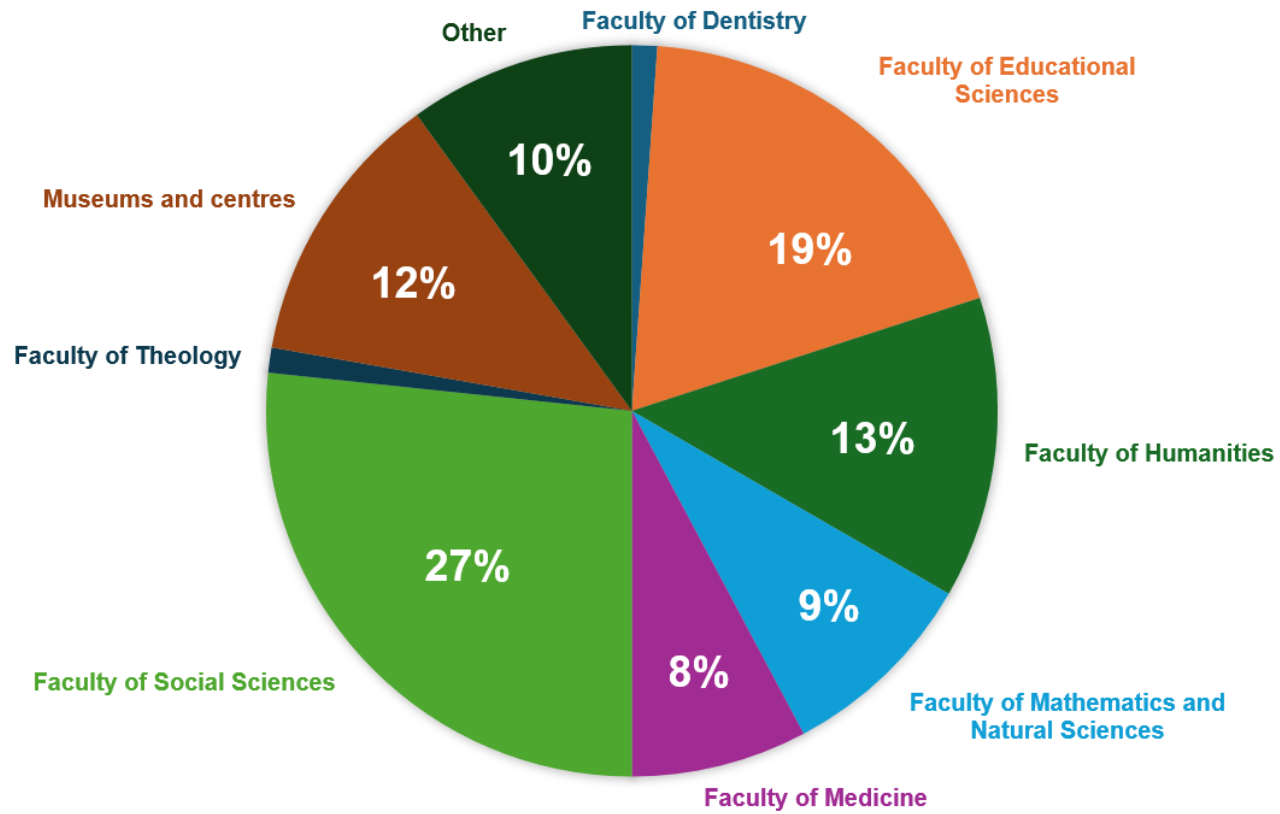
The survey is completely anonymous. The target group is all UiO-affiliated academic staff who work fully or partly with qualitative data.

By qualitative data we mean here all data that is not numerical (for example text, sound, video, image). This data is often context-sensitive, and may contain personal information or copyrighted material. Therefore, it also requires safe solutions for storage and processing. It can be also challenging to share or reuse such data.

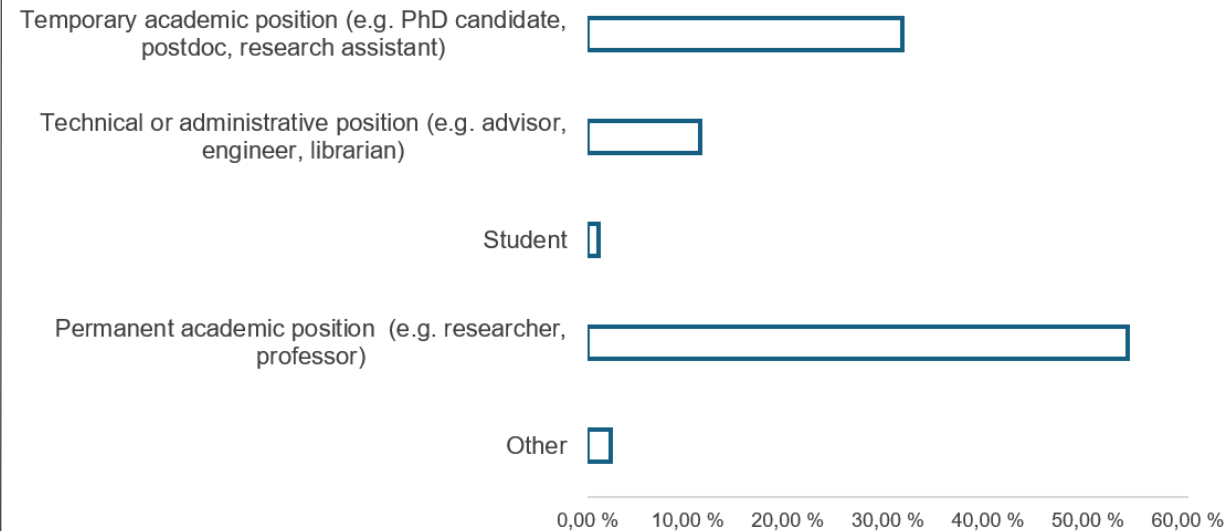
The survey is a part of [QualiFAIR's](#) work on infrastructure for sharing and reuse of qualitative data. If you have any questions about the survey, please contact QualiFAIR's coordinator Agata Bochynska via email: agata.bochynska@ub.uio.no

N = 93

Norwegian version N = 68
English version N = 25



What is your current position?



Type, formats and classification of data

Type, formats and classification of data

'Everybody' collects or generates **new (primary) data**, but **secondary data** is also used in about 40%.

Type, formats and classification of data

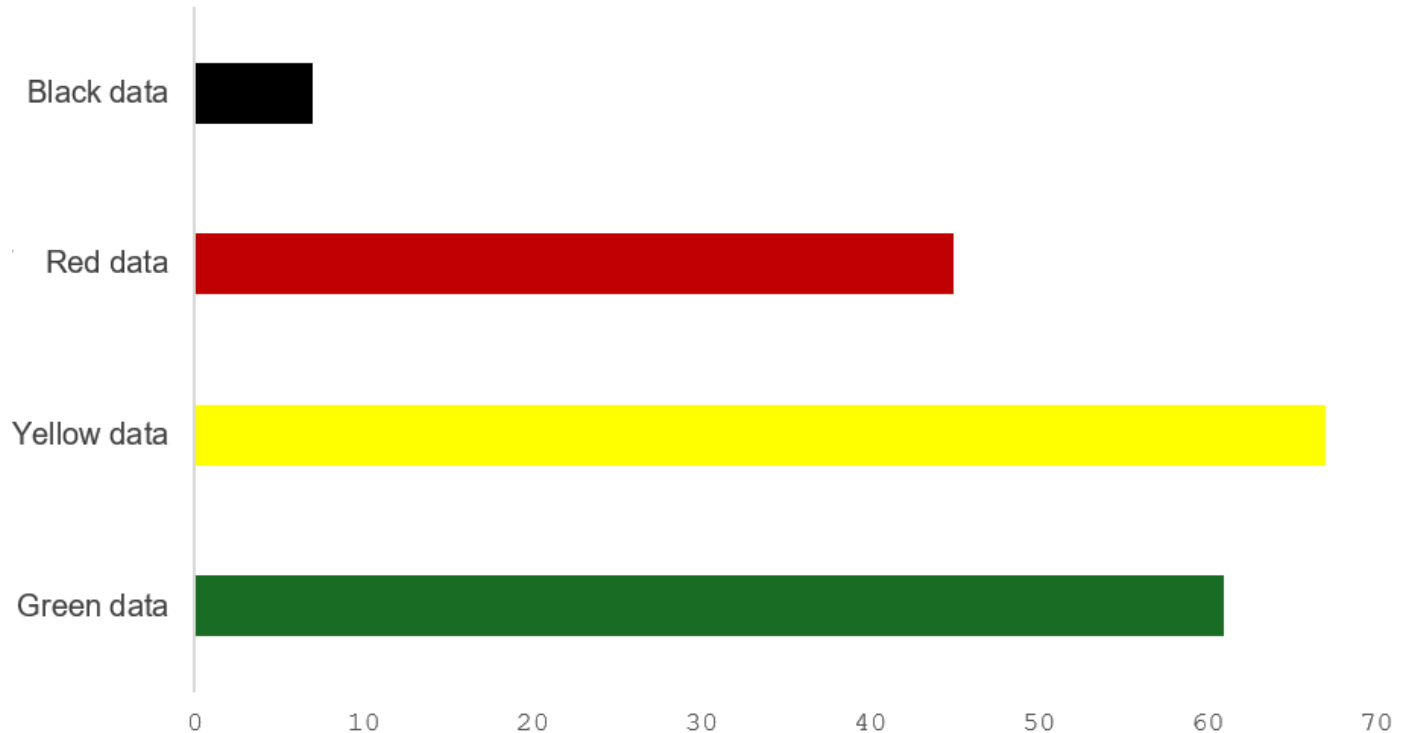
‘Everybody’ collects or generates **new (primary) data**, but **secondary data** is also used in about 40%.

- Data types predominantly consist of **text** at approximately 80%
- **Audio** around 65%
- **Images** and **videos** at about 50%
- with a small percentage encompassing **other data formats**.
These other data formats are most commonly numerical data (such as sensor data, brain scans, climate, 3D, and other test data) or text in scanned images, PDFs, archives, and large databases

Type, formats and classification of data

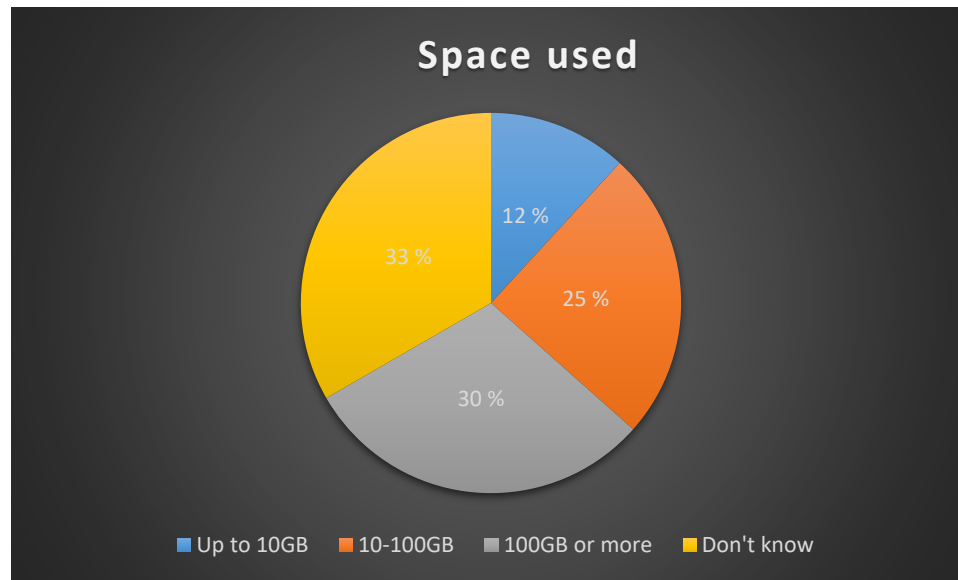
Regarding the classification of the datasets in question, approximately

- 65% are denoted as green data,
- Three-quarters as yellow data
- 50% as red data
- less than one-tenth, as black data.



During the project:
data storage

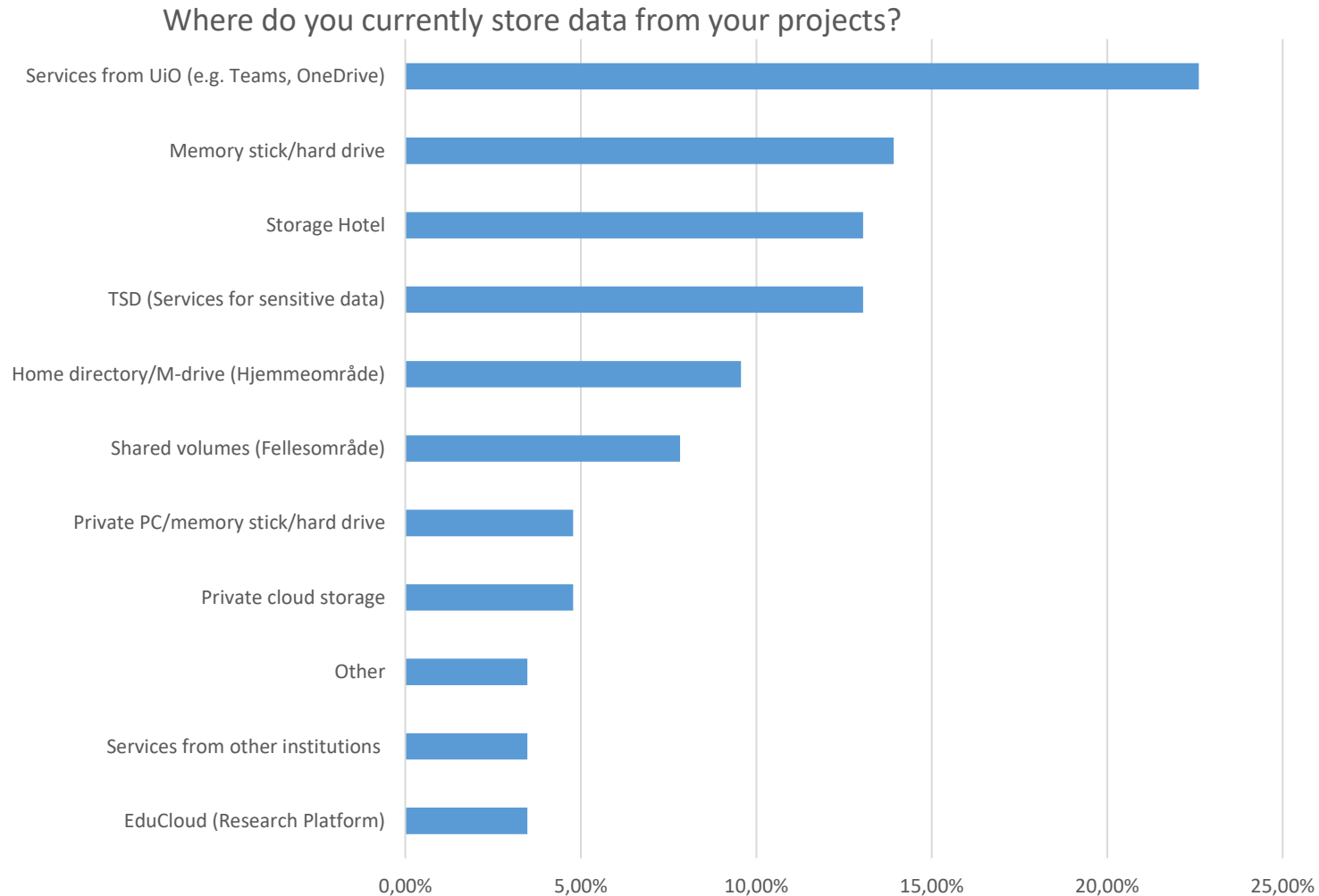
How much space do you need to store the data? Do you pay for storage space?



There are minor differences in responses between Norwegian and English versions of the survey concerning who bears the **cost of storage**. In the majority of instances, it is either the **institution** or the **funding body** that covers the expenses, and in only a few cases does the researcher personally assume the financial responsibility.

The need for storage is not fully met, partly because new data is continuously being generated and partly do to need for easy sharing with external researchers.

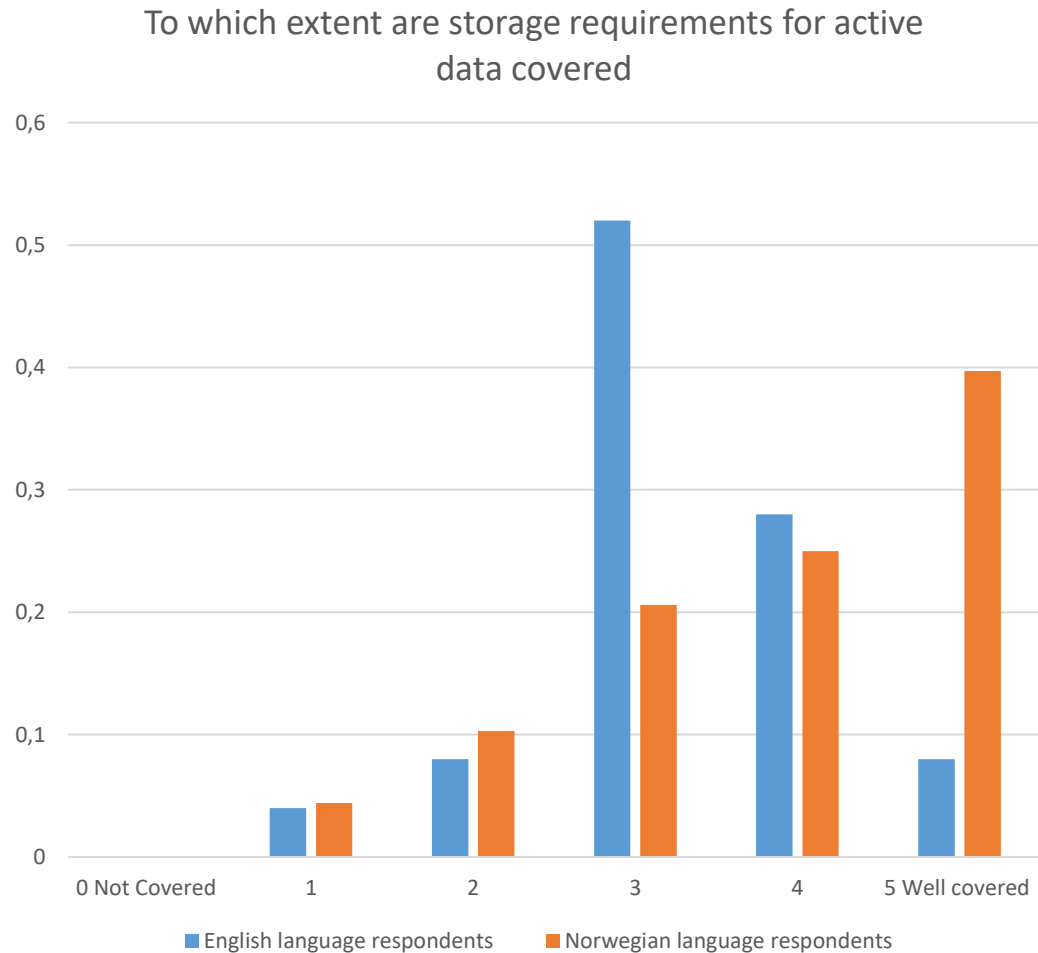
Data storage during the project



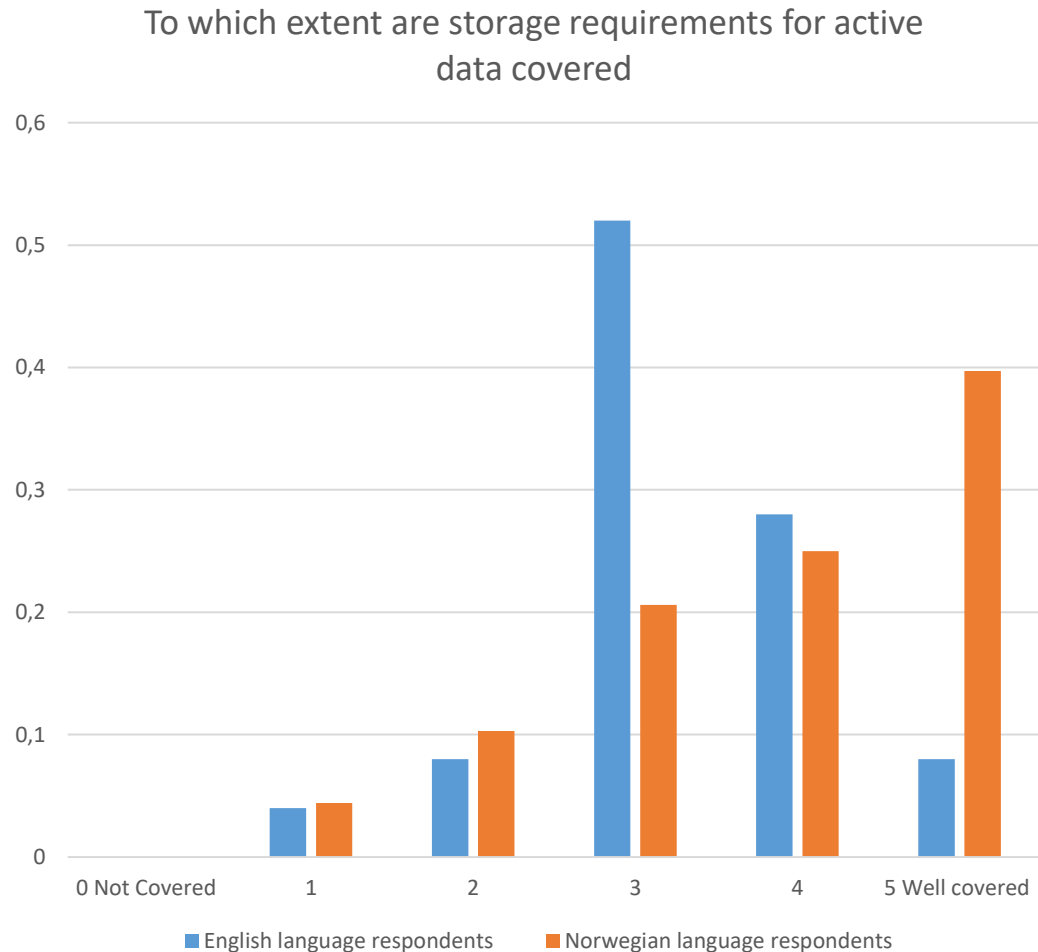
There are some differences in the responses between those who completed the questionnaire in Norwegian and those who opted to fill it out in English. In general, those who filled out the survey in Norwegian appear to utilize more of UiO's (University of Oslo's) own services, such as TSD (Services for Sensitive Data), storage hotel and the like, while those completing it in English tend to make less use of UiO's services and more use of private equipment and cloud solutions. The use of UiO Teams and OneDrive is relatively similar between the two groups.

Other options for data storage encompass platforms like the UiO Dropbox, the National Infrastructure for Research Data (NIRD), custom local arrangements, and other cloud services

To what extent are your storage requirements (for 'active' data while the project is ongoing) covered?



To what extent are your storage requirements (for 'active' data while the project is ongoing) covered?



The University of Oslo's (UiO) **websites** are the preferred source for information on storage and data processing solutions at UiO, followed by inquiries to **local IT or laboratory personnel**, and in third place, **collaboration partners or colleagues at UiO**. Additionally, the IT department at UiO (USIT), **research advisors** at the department / faculty and supervisors are also pointed to as possible sources of more information about storage and processing possibilities at UiO

‘What works well or what is still missing in the storage solutions you use now?’ - This was an open-ended question and the responses were numerous and, to some extent, divergent.

Happy users say:

Good, storage hotel works well,
nettskjema and apps works well,
UiO Onedrive works well, TSD
and storage hotel works well
when you have learned to use it,
Educloud works well, everything
is okay for the data types I use,
and many users are satisfied

‘What works well or what is still missing in the storage solutions you use now?’ - This was an open-ended question and the responses were numerous and, to some extent, divergent.

Happy users say:

Good, storage hotel works well, nettskjema and apps works well, UiO Onedrive works well, TSD and storage hotel works well when you have learned to use it, Educloud works well, everything is okay for the data types I use, and many users are satisfied

Not so happy users say:

Local data handling challenges (space restrictions in different forms). Lacks of an adequate system for retrieving and sharing data. Concerns about scalability with increased data generation, cost, challenges with local encrypted drives. Wishes TSD was easier to use, easier to share data with international researchers / collaborators, Teams and sharing data outside of UiO runs into trouble and could really need an easy to use cloud solution for red data and so on

‘What works well or what is still missing in the storage solutions you use now?’ - This was an open-ended question and the responses were numerous and, to some extent, divergent.

Happy users say:

Good, storage hotel works well, nettskjema and apps works well, UiO Onedrive works well, TSD and storage hotel works well when you have learned to use it, Educloud works well, everything is okay for the data types I use, and many users are satisfied

Not so happy users say:

Local data handling challenges (space restrictions in different forms). Lacks of an adequate system for retrieving and sharing data. Concerns about scalability with increased data generation, cost, challenges with local encrypted drives. Wishes TSD was easier to use, easier to share data with international researchers / collaborators, Teams and sharing data outside of UiO runs into trouble and could really need an easy to use cloud solution for red data and so on

During the project:
data processing

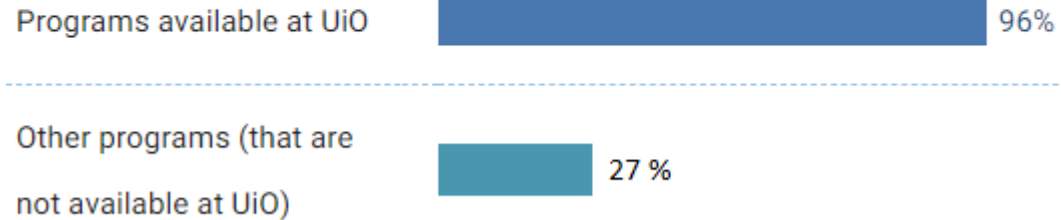
Overview of Software Usage



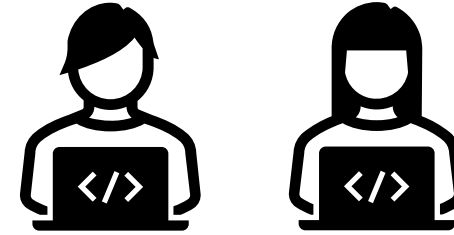
We asked about:

- 1) Which programs do you use?
- 2) What is missing?
- 3) What works well?

When working with data (processing and/or analyzing), what kind of programs do you use?



Overview of Software Usage



UiO Software used:

- **NVivo:** 31 mentions
- **Excel:** 7 mentions
- **Office Software:** 6 mentions
- **Word, R, Adobe Photoshop, Python, Autotext:** 4 mentions each
- **SPSS, Adobe Illustrator:** 3 mentions each

Software Not Available at UiO:

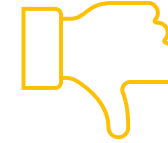
- **3D Modeling and Visualization:** Blender, MeshLab, Reality Capture, Metashape
- **Qualitative Data Analysis:** NVivo, ATLAS.ti, Taguette (for Linux)
- **Image and Video Editing:** (Adobe Premiere Pro, Adobe Acrobat Pro), Imaris
- **Document and Reference Management:** Zotero, Obsidian (for note-taking and data organization)
- **Transcription and Text Analysis:** Olympus transcription system with foot pedal, Gephi, MaxQDA
- **Photography and Color Processing:** Rawtherapee, Basic Colour
- **Data Cleaning and Management:** Open Refine, Custom software (often developed in-house for specific needs)

Feedback on Software Performance and Improvement Needs



What Works Well:

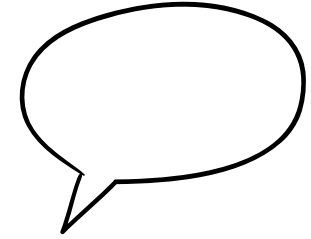
- **Installation & Support:** Effective installation on UiO machines; good ongoing support.
- **Accessibility & Training:** Positive feedback on software availability and training options.



Areas for Improvement:

- **NVivo Licensing and Functionality Issues:** License restrictions, need for better collaboration support. NVivo cloud wanted . NVivo on TSD problems.
- **User Experience:** Difficulties with TSD platform, need for easier access.
- **Updates & Version Management:** Challenges in managing updates across multiple machines.

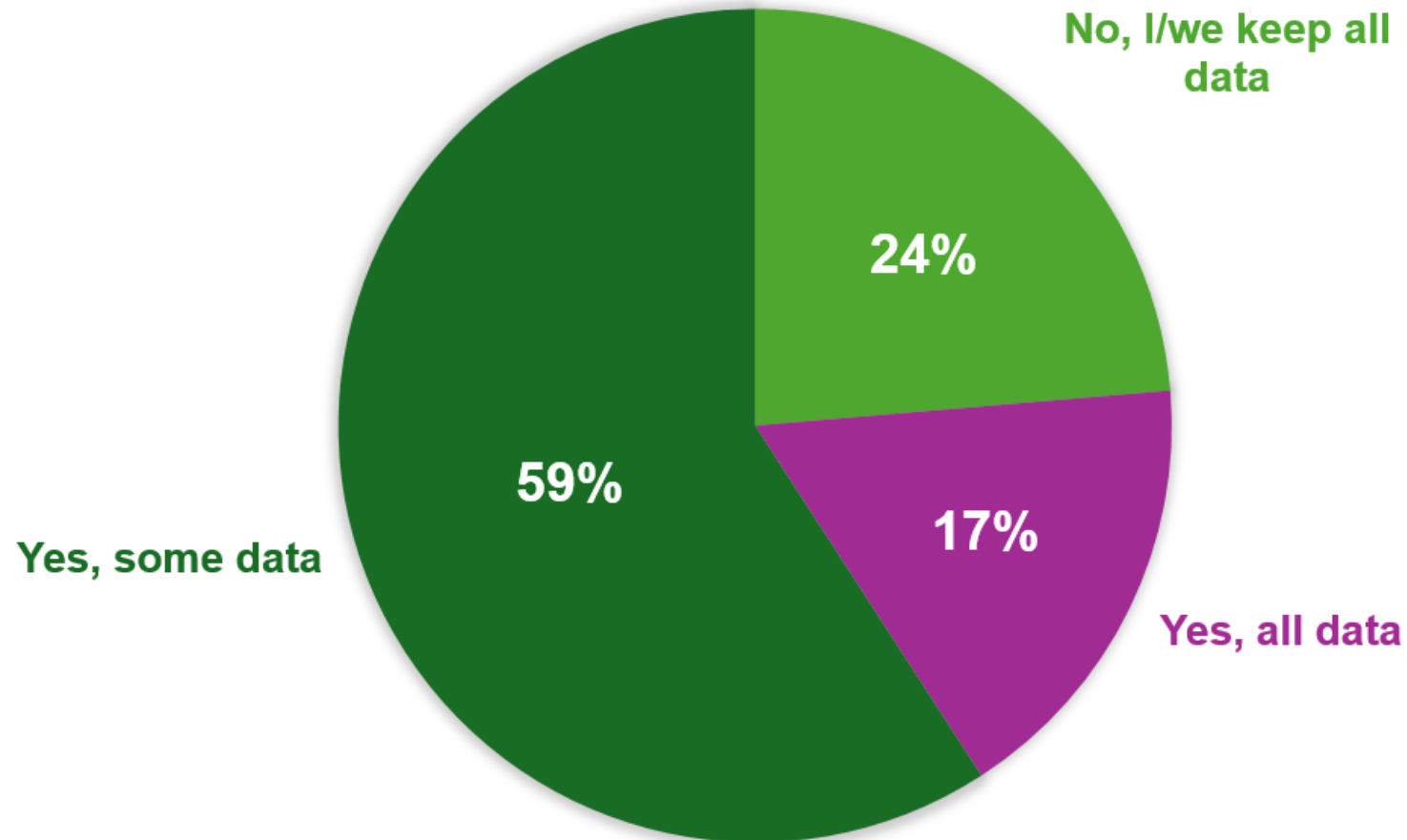
Detailed Feedback from Users



- 1. NVivo Licensing Issues:** "Sharing data in NVivo during coding is not possible without a license. This could be something to consider at UiO centrally, as it cannot always be covered by the research groups."
- 2. Technical Issues with NVivo:** "I have had very bad experiences with NVivo, which has several times crashed or deleted my latest analyses. I receive very good support from the IT department at the institute, but when it comes to NVivo, it has not been possible to get sufficient help."
- 3. Complexity of TSD Usage:** "TSD is difficult to use, and it is cumbersome (and expensive?) that students must have their own TSD folder for thesis and master's projects. I wish there was a common folder (but with limited access to each individual's project folder). This takes a lot of time for the students, or the advisors if they are the ones who have to create the TSD folder."
- 4. Desire for Better Software:** "I would like to have Interact available at UiO. We buy a certain number of licenses at the institute/project, and it is expensive and somewhat limiting."

After the project:
long-term storage and
archiving

Do you **delete** data after the end of the project?



Ethics and legal regulations

No permission for other uses of data

Worried about re-identification

Afraid of making a mistake

Where do you **store** or **archive** the data after the project is finished?

- **TSD: 1**
- **Educloud: 1**
- **UiO-servers/local storage: 12**
- **UiO Cloud storage (e.g. OneDrive): 3**
- **Private cloud storage: 1**
- **External hard drive: 3**
- **Private hard drive: 2**
- **Duo: 1**

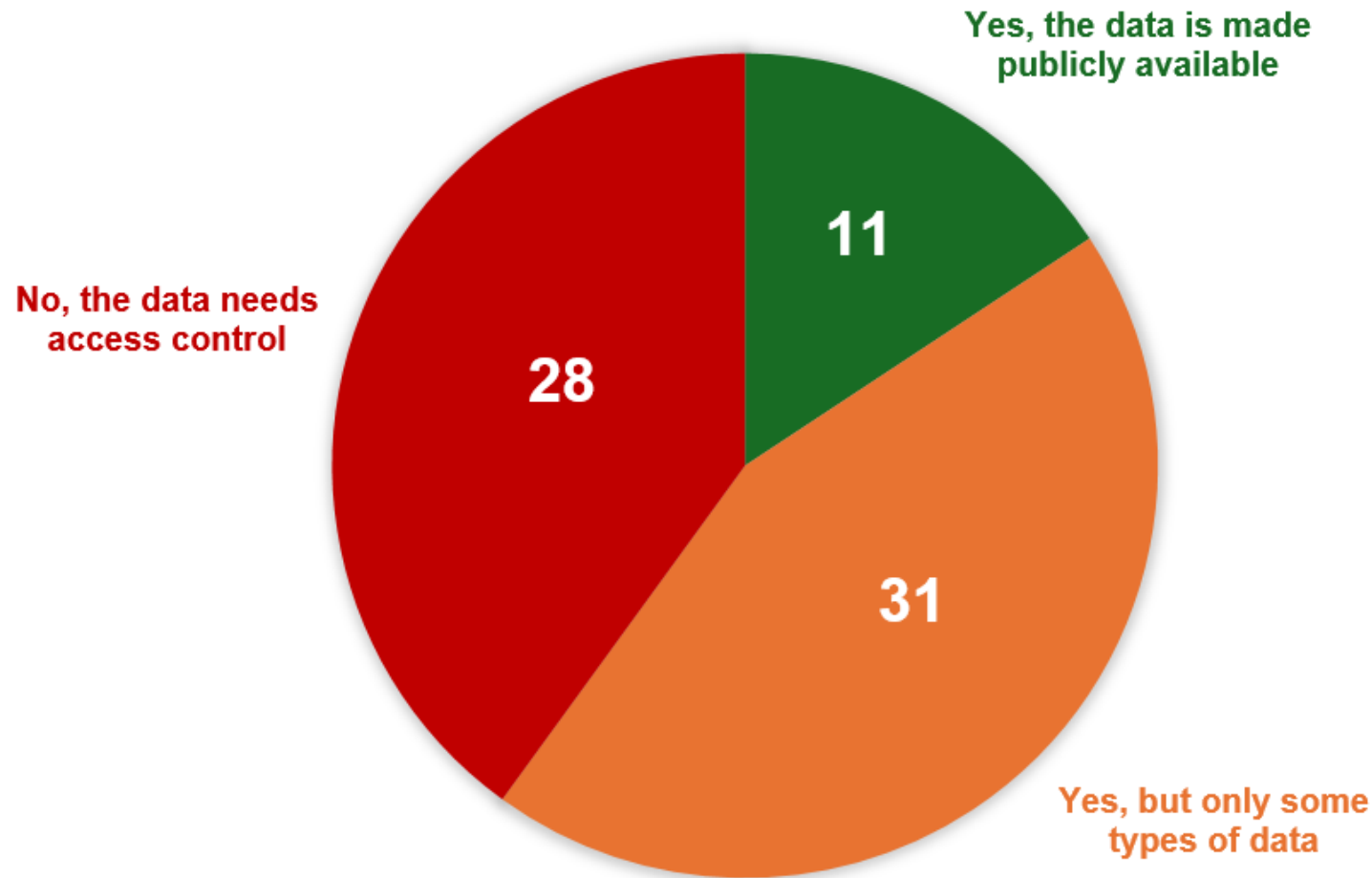
- **UNIMUS: 3**
- **Musit: 1**
- **Intrasis: 1**
- **ADED: 1**

- NIRD: 1**
- Zenodo: 1**
- Open Science Framework (OSF): 1**
- DataverseNO: 1**
- Sikt: 1**

- GitHub: 1**

Do you make data **publicly available**?

(e.g. archive or publish openly)



Do you **license** the data that you publish or archive?

Data:

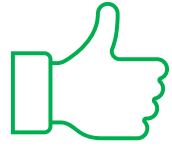
- Creative Commons Attribution (CC-BY): 3
- CC BY Share-Alike: 2
- CC BY Non-Commercial: 2
- General Creative Commons: 2

Code:

- MIT: 1

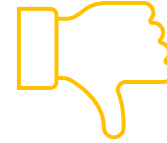
Don't use or lack of knowledge: 21

Feedback on long-term storage and improvement needs



What Works Well:

- Security and easy access
- Backups and IT-support



Areas for Improvement:

- Lack of clear guidelines; more information needed
- Access and sharing in collaborative projects (e.g. by external partners)
- Storage space, data volume and costs
- Anonymizing and minimizing personal information
- Learning new systems/platforms
- Research support in this area

Infrastructures for storing, sharing and archiving data

Where do we store and preserve qualitative data today?

Results from an UiO survey

Chief Engineer Torgeir Christiansen, Teaching Learning Video lab (TLVlab), Department of Teacher Education and School Research, UiO

Dr Anders Lunde, Oslo Special Education and Learning Lab, Department of Special Needs Education, UiO

Dr Agata Bochynska, Open Research, University of Oslo Library

What infrastructure solutions can we provide for qualitative research at UiO?

Dr Anne Schad Bergsaker, IT-department, UiO



QualiFAIR frokostseminar 14.05.2024

QualiFAIR Closing Seminar

12 September 2024, 09:00 – 15:00
Professorboligen, University of Oslo

Keynotes:

Dr Nicki Lisa Cole, Know-Center, Austria
Dr Peter Branney, University of Bradford, UK

Panel discussions:

Ethics and privacy in open qualitative research
Reuse and multiple uses of qualitative data

