

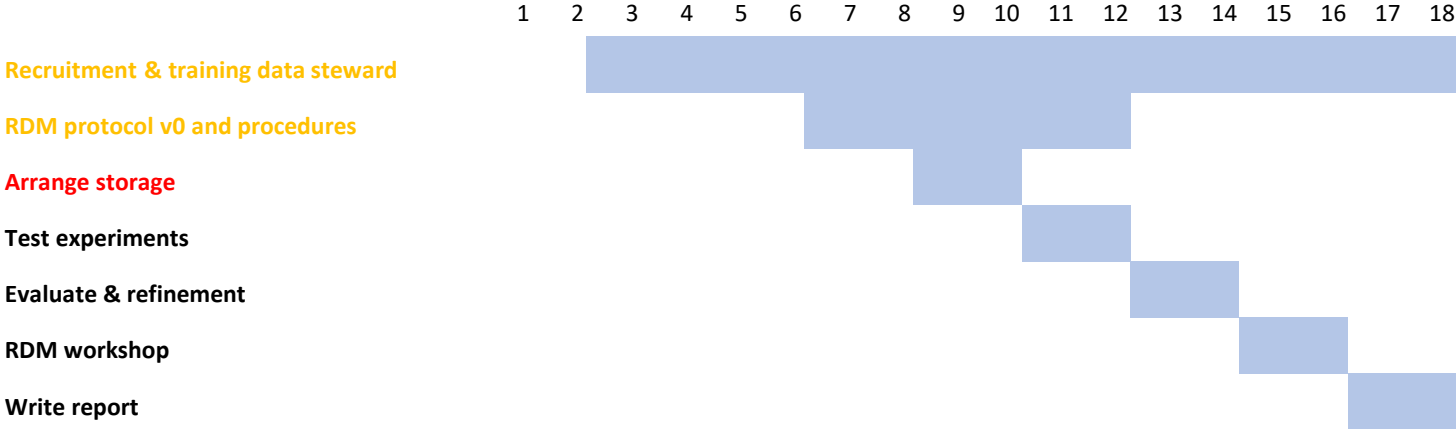


RDM at GTB

Tim van Daalen

1-2-2024

The plan



How much storage do we need?

Overview from IT

Dn Storageid	Storage Description	Dn Wuraccount	Standard quota (measure)	Quotagrante d (measure)	Over quota	Mh Usage	Usage Percent	cost per year	start	end	Costs
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-PSG_ATAL_DATA		Onbekend	20	20	20	3	13 %	€ 0.70	01-06-2023	31-08-2023	€ 3.50
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-BROCCOLIOOGSTROBOT		Onbekend	3	3	3	2	81 %	€ 0.70	01-06-2023	31-08-2023	€ 0.53
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-EFSAKASEMISSIE		Onbekend	1	1	1	1	59 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-FREEZER		Onbekend	1	1	1	0	0 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-GLASDATA		Onbekend	45	45	45	36	80 %	€ 0.70	01-06-2023	31-08-2023	€ 7.88
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-IPMETHIOPIE		Onbekend	1	1	1	0	35 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-KWALITEITSZORGPROEFBEDRUF		Onbekend	1	1	1	0	3 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-LIGHTSYM2012		Onbekend	1	1	1	0	42 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-OS MANAGEMENT		Onbekend	1	1	1	0	34 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-PARAPLUPLAN GERBERA		Onbekend	2	2	2	2	86 %	€ 0.70	01-06-2023	31-08-2023	€ 0.35
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-PLUGINSPECTIE_CULTILENE		Onbekend	2	2	2	2	86 %	€ 0.70	01-06-2023	31-08-2023	€ 0.35
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-PROGRAMMA_PRECISIETECHNOLOGIE_TUINBOUW		Onbekend	1	1	1	0	2 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-SAFE FLOWER CHAINS		Onbekend	2	2	2	1	36 %	€ 0.70	01-06-2023	31-08-2023	€ 0.35
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-SOFTWARE		Onbekend	5	5	5	4	78 %	€ 0.70	01-06-2023	31-08-2023	€ 0.88
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-TOMATO		Onbekend	1	1	1	1	62 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-TTI-VIRTUAL ROSE		Onbekend	1	1	1	1	81 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-WATERGY		Onbekend	2	2	2	1	72 %	€ 0.70	01-06-2023	31-08-2023	€ 0.35
FS01/IFS/MIXED/DBL-STANDARD/PROJEI/PROJECTS-WATERSTREAM		Onbekend	1	1	1	0	35 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PSG/P/PSG-GLASTUINBOUWPROJECTEN-112_ARCHIEFAGESLOTEN		Onbekend	275	275	275	40	15 %	€ 0.70	01-06-2023	24-07-2023	€ 28.34
FS01/IFS/MIXED/DBL-STANDARD/PSG/P/PSG-GLASTUINBOUWPROJECTEN-112_ARCHIEFAGESLOTEN		Onbekend	75	75	75	40	53 %	€ 0.70	24-07-2023	31-08-2023	€ 5.54
FS01/IFS/MIXED/DBL-STANDARD/PSG/P/PSG-GLASTUINBOUWPROJECTEN-240_FACILITARBEDRUF		Onbekend	1	1	1	0	12 %	€ 0.70	01-06-2023	31-08-2023	€ 0.18
FS01/IFS/MIXED/DBL-STANDARD/PSG/P/PSG-GLASTUINBOUWPROJECTEN-411_GEWASGEZONDHEID		Onbekend	1030	1030	1030	1029	100 %	€ 0.70	01-06-2023	23-07-2023	€ 104.14

- 60TB on project data a € 7.200,- per year
- 2TB Archive on W a € 240,- per year

Where is the rest?

- 1 spectral image is already more than 1GB



What are our options?

Data only consists of code/scripts or models ⓘ

- Yes No

During research share data/give access rights ⓘ

- No sharing
 Share within group
 Share cross-departmental
 Share beyond WUR

My dataset exceeds 1 TB ⓘ

- Yes No

Data Classification ⓘ

- Negligible
 Some
 Serious
 Disruptive

Do you have specific needs in terms of accessibility? (multiple answers possible) ⓘ

- Long term storage/archiving data (more than 10 years)
 When I leave WUR or (unexpectedly) drop out, data is still available to my research group
 Have access without my own laptop/ pc
 From a mobile device (tablet or smartphone)
 Offline access, e.g. to work on a boat

W: drive (Enterprise File Storage) WUR network drive suitable for storage of small amounts of data (<1TB) Add to compare <input type="checkbox"/>	W: drive (Massive File Storage) WUR network drive suitable for (cheap) storage of large amounts of data (>1TB) Add to compare <input type="checkbox"/>	W: drive (Massive File Storage Disaster Recovery) WUR network drive suitable for cheap and disaster proof storage of large amounts of data (>1TB) Add to compare <input type="checkbox"/>
W: drive PROJECTS WUR network drive suitable for cross-departmentally storing and sharing data Add to compare <input type="checkbox"/>	Teamsite (Sharepoint online) An Office 365 SharePoint site to store and share files Add to compare <input type="checkbox"/>	Microsoft Teams Environment to work together and store and share files Add to compare <input type="checkbox"/>
OneDrive for Business An Office 365 personal cloud service to store and share files Add to compare <input type="checkbox"/>	Yoda A data management service to deposit, share, publish and preserve research data Add to compare <input type="checkbox"/>	Git@WUR An environment to store and collaborate on source code and statistical models Add to compare <input type="checkbox"/>
M: drive A personal storage space on a WURnet account Add to compare <input type="checkbox"/>	IT-servicedesk Please contact IT support.	Information Security Officer Please contact your local Information Security Officer.

<https://library.wur.nl/storagefinder/>

Contact data@wur.nl.
or other IT services.

Yes X Serious X

W: drive (Enterprise File Storage)

WUR network drive suitable for storage of small amounts of data (<1TB)

Add to compare

W: drive (Massive File Storage)

WUR network drive suitable for (cheap) storage of large amounts of data (>1TB)

Add to compare

W: drive PROJECTS

WUR network drive suitable for cross-departmentally storing and sharing data

Add to compare

Teamsite (Sharepoint online)

An Office 365 SharePoint site to store and share files

Add to compare

OneDrive for Business

An Office 365 personal cloud service to store and share files

Add to compare

Yoda

A data management service to deposit, share, publish and preserve research data

Add to compare

M: drive

A personal storage space on a WURnet account

IT-servicedesk

Please contact IT support.

Yoda



Description

Yoda is a research data management service that enables WUR researchers and their partners to securely deposit, share, archive and publish large amounts of research data during all stages of a research project. Yoda allows users to add metadata at the folder level. Yoda is *not* meant for storing *non-research data* (e.g. invoices or project contracts). If you have already been added as a Yoda user, go to <https://yoda.wur.nl>. If you want to start using Yoda, please contact your data steward or data@wur.nl

Example use

Collaborate with internal and external parties on research data. Transfer data from and to Yoda (via an API), for instance research outputs from eLab to Yoda. You can get a Yoda-environment on group level (e.g. chair group or business unit) or project level (e.g. project team).

Costs

Free, on condition of the [Yoda@WUR Terms of Use](#). For example, when you transfer >5TB of data to Yoda, contact data@wur.nl beforehand.

(Max.) Storage capacity

Unlimited. Please contact data@wur.nl if you intend to expand storage capacity or if storage space exceeds 25 TB.

Data classification

Negligible, Some, Serious (If you have questions about data classification and/ or security, contact your local [Information Security Officer](#))

Data sharing

Yes, files can be shared internally as well as externally.

Access rights

Yes, a data manager in the chair group/ business unit or project group can invite (external) users (to a folder)

ta

Yoda free?

Service	Price per TB per year
Yoda hosting	€ 26.000,-
CEPH storage + 2 Backups	€ 246,24 + € 305,28 = € 551,52
Object storage	€ 152,64

Our prediction:

- The BU will produce around 60TB/year
- We have to store the data for 10 years
- Over 10 years at least:
 - CEPH: $60 * 551,52 = € 33.091,2/\text{year}$
 - $9*60 = 540\text{TB}$ object storage: € 82.425,6/year
 - Total: € 115.516,8

Our BU manager is only willing to invest
€ 20.000 – € 30.000...

W with disaster recovery € 120,-/TB/year: $600 * 120 = € 72.000/\text{year}$

And now?

- iRODS?

SURF**

Service	Price per TB per year
iRODS hosting	€ 11.720,-
CEPH storage	€ 215,28
Object storage	€ 152,64
Tape storage*	€ 12,32

- 60 TB object storage: € 9.158,4/year
- 540 TB Tape: € 6.652,8/year
- Total: € 15.811,2

WUR

Service	Price per TB per year
iRODS hosting	€ 23.320,-
File storage with DR	€ 120,-
Tape storage with duplication	€ 20,-

- 60 TB file storage: € 7.200/year
- 540 TB Tape: € 10.800/year
- Total: € 18.000,-

Costs are not everything:

- Reliability
- Support

* Paid in 5-year terms, **no duplication!**

** SURF defined fixed storage types for YODA but gives complete freedom with iRODS.

iRODS!

- Lot to be done:
 - Graphical user interface
 - Access
 - Robustness
 - Moving data from disk to tape
 - A coupling with HPC and other services
 - ...

Note to the director to ask for funds



Next steps

- Trying iRODS
- Creating overview:
 - For the researcher:
 - An A4 with simple stepwise procedures
 - For the Data steward:
 - Which projects have started?
 - Do they have a data management plan?
 - Which projects have finished?
 - Data, reports?
 - For the BU manager:
 - Costs
 - How often is data reused/retrieved

Inventorize key points

- Metadata?
 - How to find the data again
 - Make sure we can (re)use it
- Connecting our facilities?
 - Generating FAIR data is easier than motivating researchers to make it FAIR