Meeting date: 13 March 08:30 – 10:00
Meeting place: Teams
Meeting subject: Meeting # 2 – 2022, AGRO Data Management Committee

Attendees: René Gislum, Enoch Narh Kudjordjie, Maria Knadel, Mette Vestergaard Odgaard, Jens Grønbech Hansen, Christina Rønn Ingvardsen, Anders Almskou-Dahlgaard, Lea Kjærgaard Eriksen, Jytte Christensen

Agenda:

1. Approval of agenda
2. Messages:
    a. New member of the committee (Lea Kjærgaard Eriksen)
    b. Activities at Faculty and National level / Jens
    c. Other messages from committee members
3. Follow up on recommendations from last meeting and related to the tasks of the committee/ all
    a. Repositories, - possibilities and roadmap / Jens and Maria
    b. Opportunities to develop a proof of concept 'Field Trial Planning Tool' / Jens BK & Jens GH
    c. Use of videos for our procedures and protocols / Jens
    d. Response from Institute leadership on our recommendations from last meeting
    e. How to get "flesh on the bones", priorities and recommendations
4. E-Doc repository – searchable / Christina and Karin
5. Committee website. What else to put there?
6. AOB
    a. Next meeting 29 August 2022 (new meeting date to be decided)

Ad 1: Approval of agenda
Approved

Ad 2: Messages
Ad 2a: Lea Kjærgaard Eriksen is new member of the data management committee. She is employed as project coordinator and fundraiser and she is based at Flakkebjerg.
Lea will assist Jens on seeking external funding for helping to fulfil our aim for datamanagement. One of her tasks is to help implement data management procedures e.g. data from published articles.

Climate and Water
Aarhus University
Blichers Allé 20
PO box 50
DK-8830 Tjele
Denmark

Tel.: +45 8715 6000
E-mail: agro@au.dk
Web: agro.au.dk/en

Ad 2b:
At AU level the digital transformation has so far had its focus on administrative issues and not research itself.

ERDA and ERDA SIF implementation is postponed until September. AGRO should deliver and work on 1-2 use cases in the implementation phase and this is also postponed. The AU Open Science Forum as well as the subgroup on Data management coordination is also not yet activated (Jens is in both). It is important that AGRO comply with the AU and national strategy on data management – also to make use of new opportunities internal at AU as well as via DeiC.

Jens will try to organise a workshop at AGRO on FAIR data management with some use cases after summer. Torben Brejnegaard Nielsen and Birte Christensen-Dalsgaard are willing to visit our department. Torben can talk about ERDA and ERDA SIF and Birte Christensen Dalsgaard about AU strategy on FAIR data management. This could be at the next department meeting in September or October.

At national level DeiC has established a secretariat at DTU to support the ministry working groups on implementation of FAIR data management at the universities and related institutions. Jens is in working group (a), and one of the tasks is to monitor and come up with recommendations re data repositories and meta-data platforms.

Ad 2c: Regarding big data and drone/robot data we need a place to store raw data. AU does not expect that we can do this by ourselves – we need a responsible person and therefore we need funding.

In the future there will be cross-university repositories. Until further Flakkebjerg will continue as usual.

Maria circulated an e-mail regarding a webinar on DeiC Interactive resources 22 June.

Ad 3: Follow up on recommendations from last meeting and related to the tasks of the committee

Opportunities to develop a proof of concept 'Field Trial Planning Tool'
Jens Bonderup Kjeldsen has established contact to DIS about possibilities to develop a tool for Field Trial Planning. The price for a proof of concept project is 70,000 kr. Our vision is that the system should be maintained and operated by AGRO but still with assistance from DIS consultants.

See slides #31 - #34 in the file DMC_M2_June_2022.ppt located here: O:\Tech_AGRO\Stab\Data management udvalg\Møder\2022\Møde 2 130622\

It was decided to form a working group (Anders should be included) to go into this. Jens will contact Jens B. K. regarding further steps and also contact Jørgen regarding finances. It is important that we get a system that is easy to operate.

There was a discussion on other tools e.g. ARM and Nordic Field Trial.

Use of videos for our procedures and protocols
Jens showed a short demonstration video he had produced with his cell phone as an example. It can be useful to create small videos for procedures. It is essential to find a place to file such videos.

All agreed on that it can be very useful with video guides for most procedures. As it is now new staff get introduction by experienced staff.

It was agreed that Maria, René and Anders will try to produce a protocol including videos. Anders will try to make an example including a QR code. At the next meeting we will discuss this further.

Response from department management on our recommendations
Jens have had no response, but he plan to attend a department meeting and bring this up.

How to get "flesh on the bones", priorities and recommendations

See line M2 in slides #38 - #46 + slide #50.

A general thing among most sections is the problem to install computers on the AU network when buying equipment where the computer is inclusive. Read more here https://medarbejdere.au.dk/administration/it/koeb-it-udstyr-og-software/systemanskaffelse.

Ad 4: E-Doc repository – searchable

Jens raised this question about the need for a common strategy about protocols and guideline for the procedures in the field and in the lab. Currently, it seems not to be a problem, but we might follow up on this later when we know more about the AU strategy on FAIR data management

Ad 5: Committee website
If you have relevant links for our website https://agro.medarbejdere.au.dk/udvalg-i-agro/data-management-udvalg , you can send information to Jytte.

Ad 6: AOB

Next meeting is 5 September.

Meeting #2: AGRO Data Management Committee

23 November 2021, 09:00 – 11:00

Meeting place: Teams

Dagorden:

1. Approval of agenda
2. Messages:
    a. Strategy about the O-drive, security etc. / Jens
    b. Meta data platforms, DeiC and EUREKA / Jens
    c. Faculty strategy / Jens and Maria
3. Synthesis re. feedback from sections and from researcher days / Jens
    a. Discussion / all
4. Status from the Data management project re proof of concept, Crop Rotation Experiment / Margit, Ying and Poul
5. Recommendations from the committee on FAIR data management to the Agro Leadership (based on 2 and 3) / all
6. AOB - Mette
    a. Next meeting

Datamanagement on the o-drive

To ensure that all data obtained in published research projects including PhD and postdoc studies is safely stored and accessible for later use we need to:

- store and secure data from our published papers, so that it can be recovered
- have a practice for how this data is organised, described, structured, and documented
- follow AU's current policy for how data is accessed and possibly shared

STEP by step

1. Folders on the O-drive - O:/tech_AGRO-CROP-data
2. Create a folder with your name or AU-number
3. In this folder create a folder with the Research Project Number
    1. Subfolders like: Manuscript files, Graphics, Experimental Setup, Programs and scripts, Data files and Program files
4. Save your data here

First authors are responsible for archiving

(For further information – see the handout ” *Guide to saving data in AGRO – Step by step*”

Groupwork

How do we ensure that data is saved correctly? Input to the future procedure

5 min – Individually write down all of your ideas – one idea per post-it

10 min – group discussion – talk about the different ideas and choose the top three ideas and write them on a A5 post-it

After the group-work we will collect all the post-its, which will be used in the continued data management work

Group work at Researcher days, 3 November 2021, Kolding
Input from Camilla Brodam

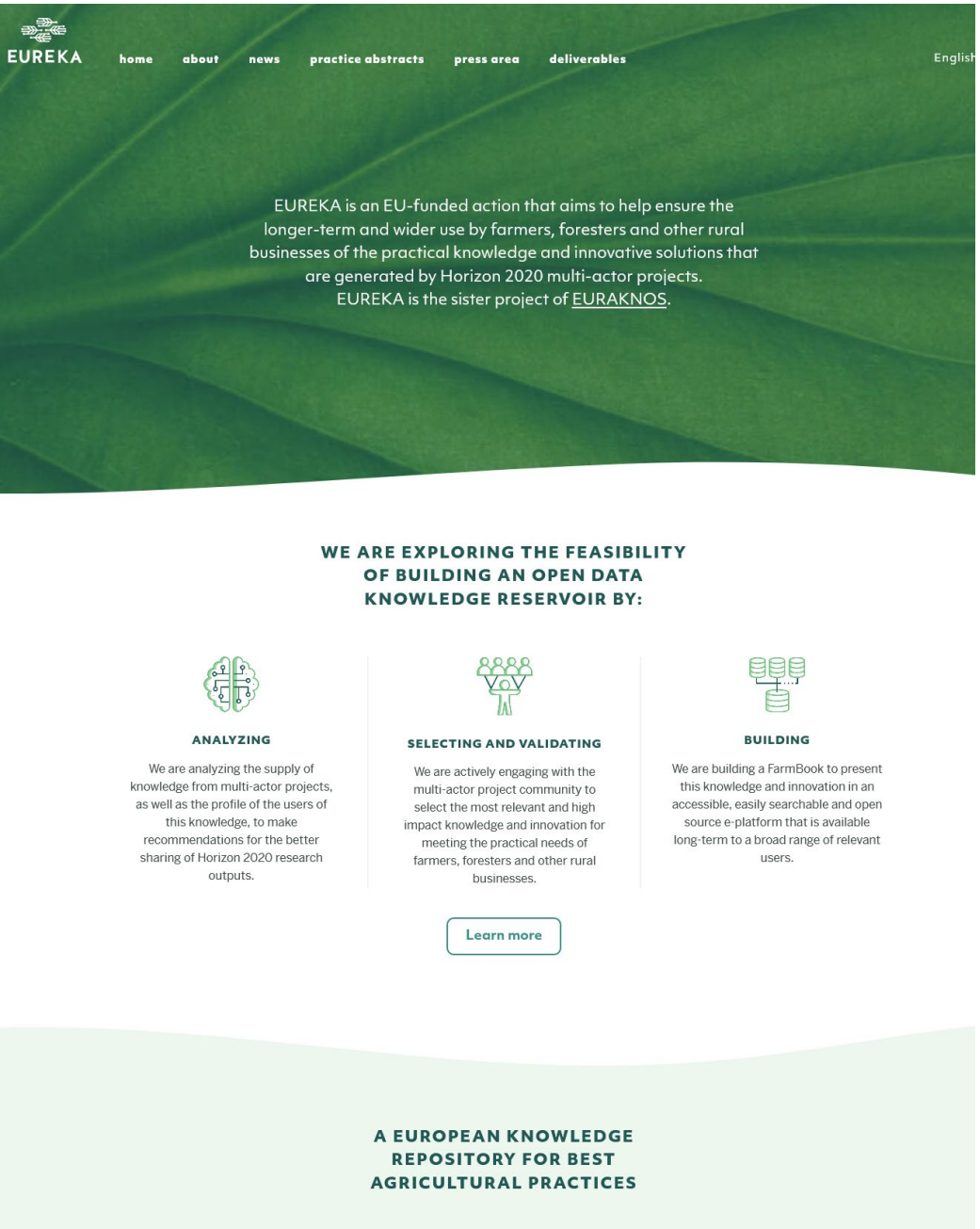
Requests and recommendations from researchers (sunthesis of inputs from group work)

- One responsible person, could be the secretary, oversee that data are stored correctly and everybody remember it. Could also be one dedicated data steward employed at the institute.
- Time allocated to manage data for everybody - ½ month per year
- Guidelines and courses how to do correctly and efficient.
- Instruct PhD students carefully. It should not be possible to submit PhD thesis before data are stored and quality controlled. The supervisor should help and oversee this.
- Responsibility for FAIR data management re Journal related data should be corresponding author.
- Include a reminder on data management every year when all PURE affiliations and metadata on papers is updated.
- Repeat again and again at section meetings, - remember to store and document all data related to scientific publications.
- Only a few expressed concerns about security on O-drive, but there was a suggestion that only a few should have access to data.

FAIR data management - M4M workshop with DeiC

A lot of homework + whole day physical 17 December 2021
– Make the wheat rust Toolbox data FAIR

Objective: Proof of concept and learn about use of Metadata



Eureka

Describe with metadata selected tools and services developed by RustWatch

A Semantic Data Model for a FAIR Digital Repository of Heterogeneous Agricultural Digital Objects

The Case of the EUREKA FarmBook

Hercules Panoutsopoulos[1], Christopher Brewster[2] and Spyros Fountas[1]

[1]*Department of Natural Resources Management and Agricultural Engineering, Agricultural University of Athens, 75 Iera Odos St, Athens, 11855, Greece*

[2]*Institute of Data Science, Maastricht University, Maastricht, Netherlands, & Data Science Group, TNO, Soesterberg, Netherlands*

Abstract

During the past few years a significant number of agriculture-related research and development projects have been implemented by receiving funding from the European Commission. All these projects have aimed to address specific problems and have produced solutions documented in the digital objects created in their context. However, the uptake of the existing information and knowledge by the concerned stakeholders is not yet adequate. On the other hand, the divergence of the available digital objects in terms of their types and formats poses significant challenges in homogeneously describing them with metadata. The Horizon 2020 EUREKA project aims to make a contribution towards this direction by developing a FAIR digital repository based on a semantic data model. The present document focusses on the design decisions related to it, as well as the rationale for the need to develop FAIR digital repositories based on formal data models.

Keywords

Agriculture, semantic data model, digital repository, FAIR principles, semantic web standards

1. Introduction

In the seven years of implementation of the Horizon 2020 Framework Programme, the European Commission invested nearly one billion euros in projects related to agriculture, forestry, and rural development[1]. These projects have created a large number of digital objects conveying valuable information and knowledge about best and innovative practices. However, the uptake and re-use of these digital objects has not yet been realised to an adequate extent. In addition, most of these digital objects are no longer available after the end of the projects. As a result, there is little potential for the various agricultural stakeholders to have access to the available knowledge for further research and solution development. This is the context in which the EUREKA

IFOW 2021: 2nd Integrated Food Ontology Workshop, held at JOWO 2021: Episode VII The Bolzano Summer of Knowledge, September 11-18, 2021, Bolzano, Italy

hpanoutsopoulos@aua.gr (H. Panoutsopoulos); christopher.brewster@maastrichtuniversity.nl (C. Brewster); sfountas@aua.gr (S. Fountas)

0000-0002-8060-9750 (H. Panoutsopoulos); 0000-0001-6594-9178 (C. Brewster); 0000-0001-9787-6268 (S. Fountas)

[1]For details, the interested reader may refer to https://ec.europa.eu/eip/agriculture/en/about/multi-actor-projects-scientists-and-farmers

Open Science ved Aarhus Universitet med fokus på Data management

DRAFT

Udarbejdet af

Birte Christensen-Dalsgaard

Indhold

Det anbefales derfor, at AU sætter sig nedenstående strategiske mål, som skal opfyldes i de kommende 5 år:

Alle forskerne på AU:

- ✓ Kender til og har forholdt sig til Open Science, herunder relevansen af FAIR – både for data og for andre af forskningens output som koder og metoder.
- ✓ Ser data management som en integreret del af forskningsprocessen og nødvendig for at sikre transparens og integritet af resultaterne.
- ✓ Ved, hvor og hvordan data håndteres og ved, hvor og hvordan man kan få støtte til data- og analysedelen af forskningsprocessen.

Forskere ved AU er aktive i forhold til:

- Udvikle praksisser og standarder for håndtering af forskningens output gennem hele forskningens livscyklus, herunder metadata standarder, standarder for data, samt hvor og hvordan disse kan gemmes under og efter projekters gennemførelse, valg af licenser og brug af persistent identifiers. (FAIRifisering)

Aarhus Universitet sikrer at:

A. Den nødvendige tekniske infrastruktur står til rådighed
B. Den nødvendige ekspertise er tilgængelig samt at der tilbydes kursus- og efteruddannelsesforløb på relevante niveauer (herunder Ph.D.)
C. Arbejdet med at dele data og andet relevant output anerkendes som forskningsrelevante aktiviteter
D. Der defineres kriterier for datas værdi ift. genanvendelighed og langtidsopbevaring og der formuleres en strategi for langtidsbevaring af data, som ikke i sin helhed er afleveret til Rigsarkivet.

Udmøntning af strategien ligger i et nyoprettet Open Science forum, hvor alle fakulteter er repræsenteret på ledelsesniveau og hvor centrale aktører i forhold til implementeringen også har sæde

Der opstilles tre scenarier for implementering på AU, som afspejler tre ambitionsniveauer for AUs på dette område.

Lavt: Her gør AU, hvad der er strengt nødvendigt i forhold til den nationale strategi for data management.

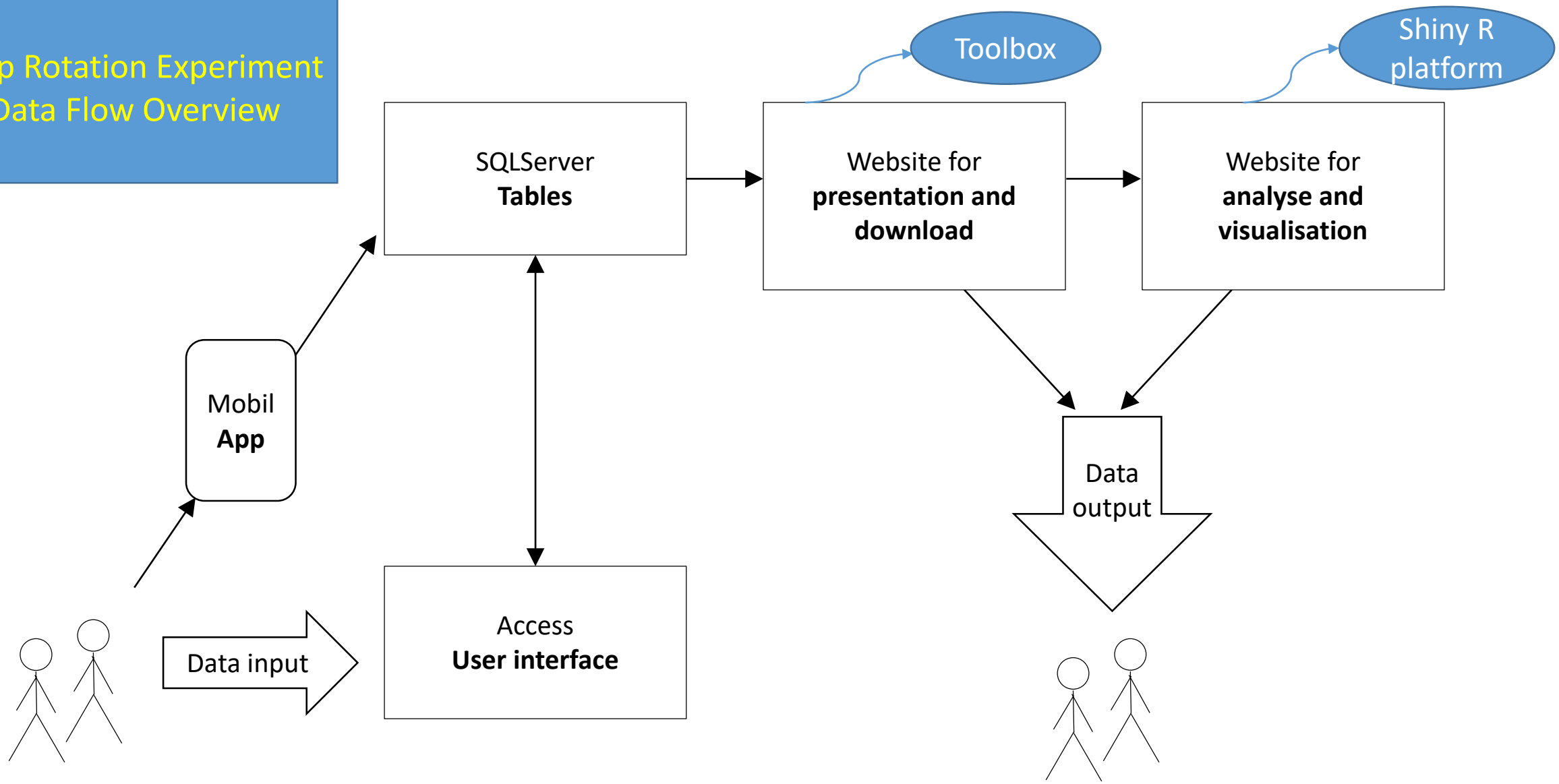
Middel: AU tager OS-agendaen alvorligt, men er begrænset i, hvor mange ressourcer, der kan bruges.

Højt: AU tager OS agendaen alvorlig og er indstillet på at gennemføre investeringer. OS-funktionen udbygges, så mere specialiserede kompetencer tilknyttes funktionen. Bemandingen på storage udvides, så AU kan være med i nationale og internationale projekter på området – sammen med interesserede forskningsgrupper. Der etableres centre/datalab på fakulteterne, som kan understøtte forskerne.

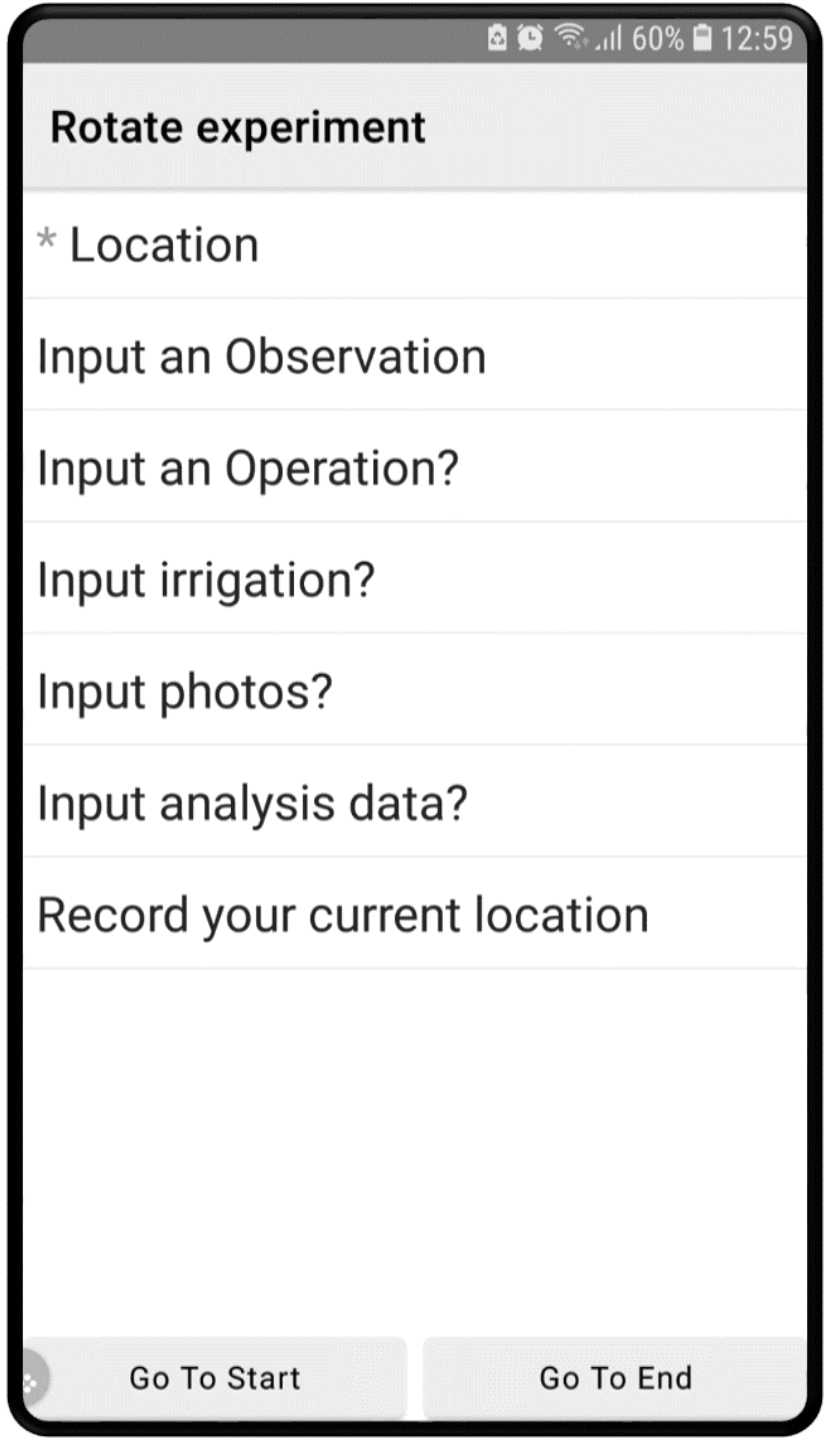
Examples of challenges and opportunities expressed by each section

Section	Challenges	Opportunities
Soil Fertility (JORNÆR)	Data storage and security on O-drive Data flow not effective	New and better equipment in the labs including automated dataflow and data management
Crop health (CROP)	Data are not safe enough using the O-Drive	Talk with ANIS as they have a nice set up for storage of big amount of data.
Entomology and Plant Pathology (PATENT)	Need a large, safe, and secure backup data storage system. O-Drive not safe enough Data on personal drives without secured and safe backup storage are at the risk of getting lost.	Data needs to be protected when stored at rest, thus, we gladly welcome the departments common data management platforms (Wheat Rust and Potato late blight Toolboxes), which is based on the "FAIR" principles.
Soil Physics and Hydropedology (JORD)	Different versions of same files and no metadata available Data flow and storage from Automated systems (drones, robots, lab equipment with automatic measurements)	The Data Management Committee should provide guidelines for metadata generation, how data should be stored and arranged
Crop Genetics and Biotechnology (CGB)	Need a platform to store sequencing data from bigger sequencing projects	Data management plan, Data stewards and an effective data management system.
Climate and Water (KLIMA)	Different versions of Access databases for same trials, difficult to extract and understand the raw data, Automated dataflow difficult.	Make a proof of concept converting the Crop Rotation Experiment data into a SQL database and including features and tools for effective data flow, data management, display and analysis of the data
Agricultural Systems and Sustainability (SYSTEM)	We do not store data on a common (open) drive. Data are located locally or in a messy project folder on O:\ or published in DCA/DCE reports, journals, etc.	Huge willingness to create a metadata-description of existing and future data, to clarify who to go to for more information and to increase interdisciplinary work.

Crop Rotation Experiment
Data Flow Overview



Mobile App Overview



Toolbox Overview

1. Filter data

AGRO DATA MANAGEMENT

Home | Rotate | Member

Welcome Ying Wang logout

Under construction

Data from rotate trial

Add_Yield_PA | Year 2020, 2019, 2018 | Location Foulum

Download

Create

2. Download data

Download File

Create

3. Analysis data

Data from rotate trial

Add_Yield_PA | Year 2020, 2019, 2018 | Location Foulum

Year	Location	Plot Key	Crop name	Undersown crop	First Crop	Block	Sub Block	Rotation	Cover crop	Fertilizer	Crop	Date	DM Yield	Gr
2018	Foulum	2510	Grass-clover	None	1	2	1	O2	Without cover crop	With fertilizer/manure	KG	22-05-2018 00:00:00	448,879769480226	
2018	Foulum	2360	Grass-clover	None	1	2	2	O2	With cover crop	With fertilizer/manure	KG	22-05-2018 00:00:00	264,381112744045	
2018	Foulum	2240	Grass-clover	None	1	1	2	O2	With cover crop	With fertilizer/manure	KG	22-05-2018 00:00:00	380,130072523413	
2018	Foulum	2100	Grass-clover	None	1	1	1	O2	Without cover crop	With fertilizer/manure	KG	22-05-2018 00:00:00	494,311386096153	
2018	Foulum	2510	Grass-clover	None	1	2	1	O2	Without cover crop	With fertilizer/manure	KG	22-06-2018 00:00:00	189,244546426677	
2018	Foulum	2360	Grass-clover	None	1	2	2	O2	With cover crop	With fertilizer/manure	KG	22-06-2018 00:00:00	130,679700494631	
2018	Foulum	2240	Grass-clover	None	1	1	2	O2	With cover crop	With fertilizer/manure	KG	22-06-2018 00:00:00	217,523021576681	
2018	Foulum	2100	Grass-clover	None	1	1	1	O2	Without cover crop	With fertilizer/manure	KG	22-06-2018 00:00:00	270,120675385169	
2018	Foulum	2510	Grass-clover	None	1	2	1	O2	Without cover crop	With fertilizer/manure	KG	14-08-2018 00:00:00	445,539979442995	
2018	Foulum	2360	Grass-clover	None	1	2	2	O2	With cover crop	With fertilizer/manure	KG	14-08-2018 00:00:00	333,317816497791	
2018	Foulum	2240	Grass-clover	None	1	1	2	O2	With cover crop	With fertilizer/manure	KG	14-08-2018 00:00:00	360,173358731958	
2018	Foulum	2100	Grass-clover	None	1	1	1	O2	Without cover crop	With fertilizer/manure	KG	14-08-2018 00:00:00	428,087037360796	
2018	Foulum	2640	S.barley	Ryegrass	1	2	1	O4	With cover crop	With fertilizer/manure	BY	16-08-2018 00:00:00	518,740964998408	
2018	Foulum	2580	S.barley	Grass-clover	1	2	1	O2	With cover crop	Without fertilizer/manure	BY	16-08-2018 00:00:00	282,979113736407	
2018	Foulum	2500	S.barley	None	1	2	1	C4	Without cover crop	With fertilizer/manure	BY	16-08-2018 00:00:00	616,175840769728	
2018	Foulum	2630	S.barley	Grass-clover	1	2	1	O2	Without cover crop	With fertilizer/manure	BY	16-08-2018 00:00:00	488,23079931726	
2018	Foulum	2460	S.barley	Ryegrass	1	2	2	O4	With cover crop	Without fertilizer/manure	BY	16-08-2018 00:00:00	450,381172844533	
2018	Foulum	2420	S.barley	Ryegrass	1	2	2	C4	With cover crop	With fertilizer/manure	BY	16-08-2018 00:00:00	513,790527881836	
2018	Foulum	2260	S.barley	None	1	1	2	O4	Without cover crop	With fertilizer/manure	BY	16-08-2018 00:00:00	433,36591964062	
2018	Foulum	2220	S.barley	Ryegrass	1	1	2	O4	With cover crop	Without fertilizer/manure	BY	16-08-2018 00:00:00	503,553341720288	

1 2 — Page 1 of 2, items 1 to 250 of 400.

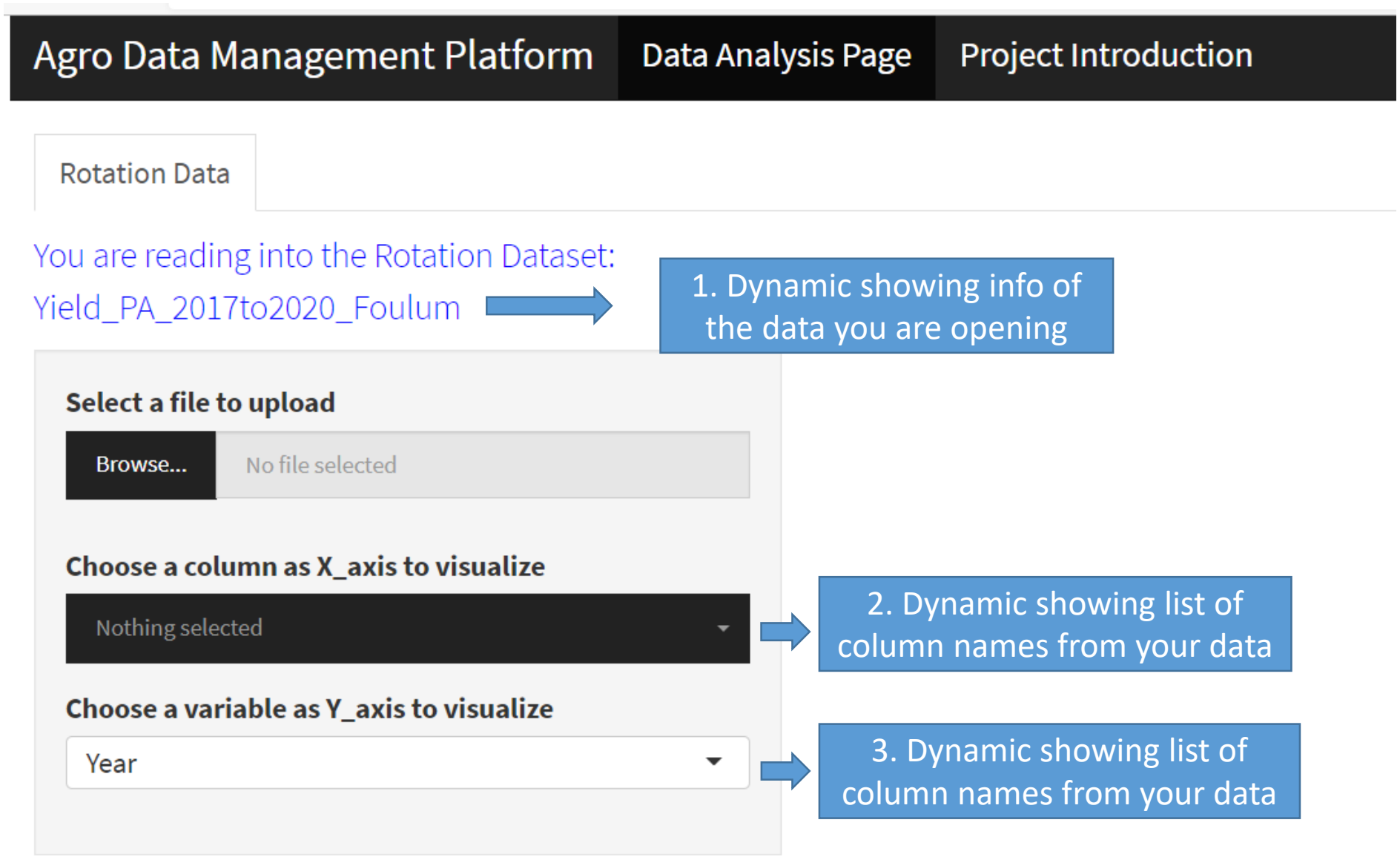
Download

Create

File

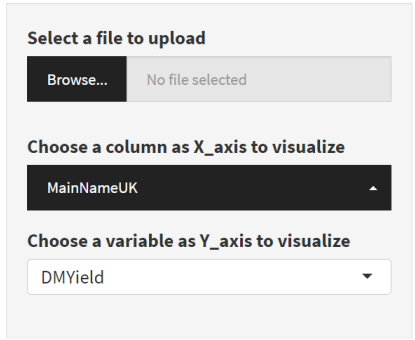
ShinyR

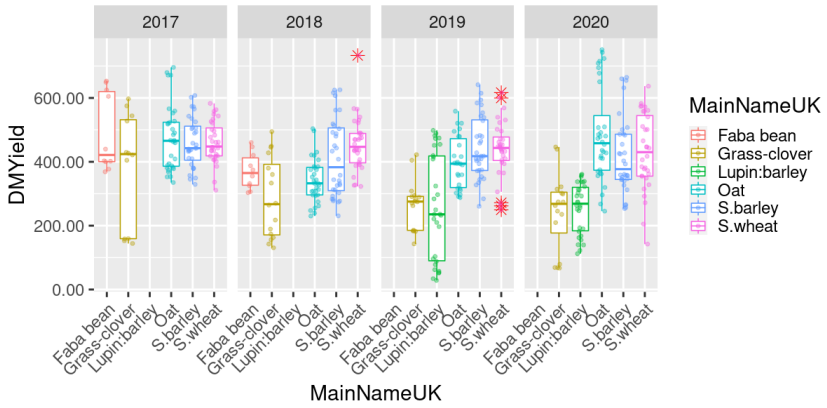
Shiny R Overview

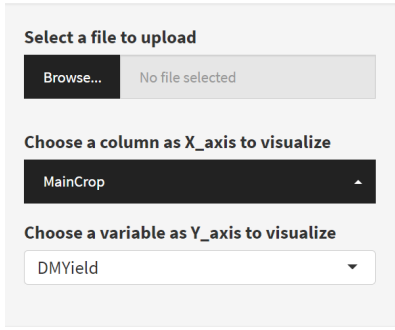


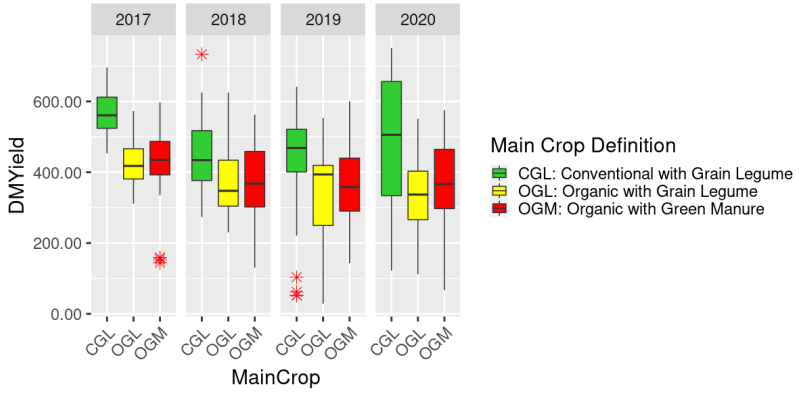
Shiny R Overview

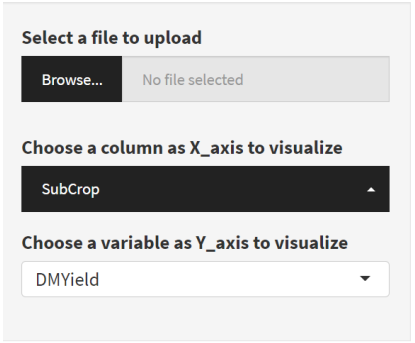
4. Demo visualizations

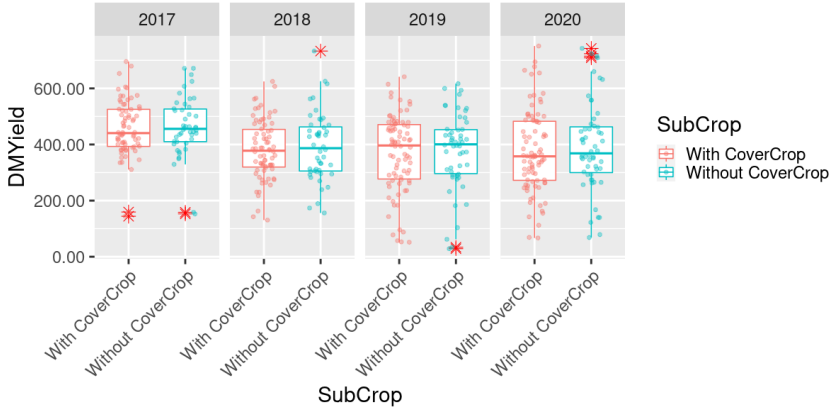


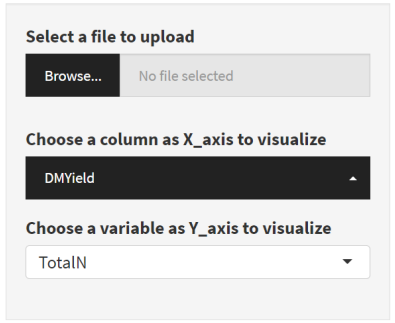


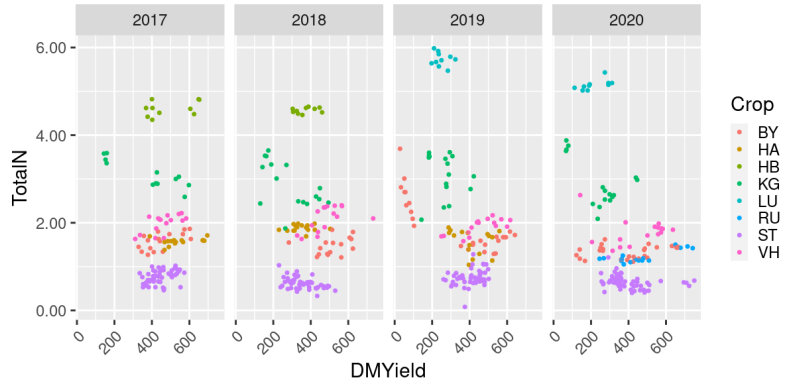












Shiny R Overview

Agro Data Management Platform | Data Analysis Page | Project Introduction

Rotation Data

You are reading into the Rotation Dataset:
Yield_PA_AllYears_AllLoctions

Select a file to upload

Browse... Yield_PA_AllYears_AllLoctions.csv

Upload complete

Choose a column as X_axis to visualize

Nothing selected

Choose a variable as Y_axis to visualize

Year

5. Upload your own dataset

Agro Data Management Platform | Data Analysis Page | Project Introduction

Rotation Data

You are reading into the Rotation Dataset:
Rotate_601442_30favgxr

Select a file to upload

Browse... No file selected

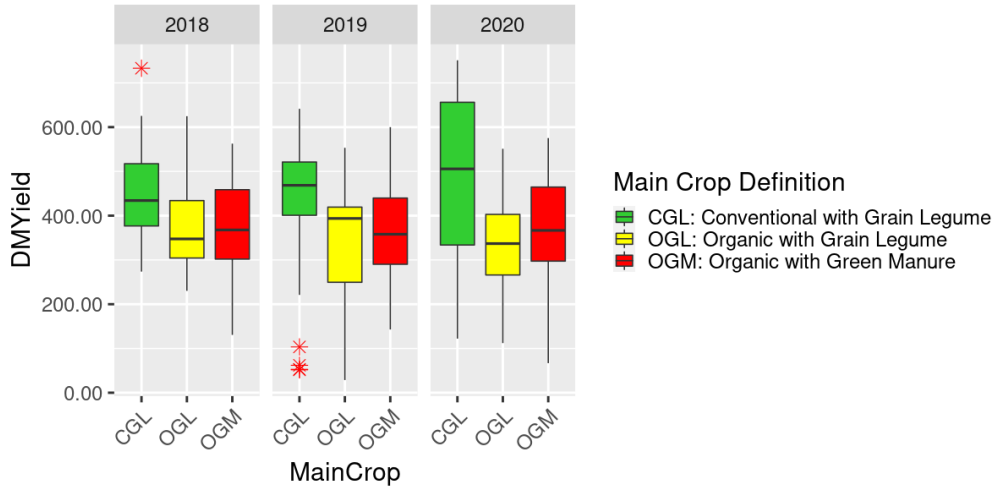
Choose a column as X_axis to visualize

MainCrop

Choose a variable as Y_axis to visualize

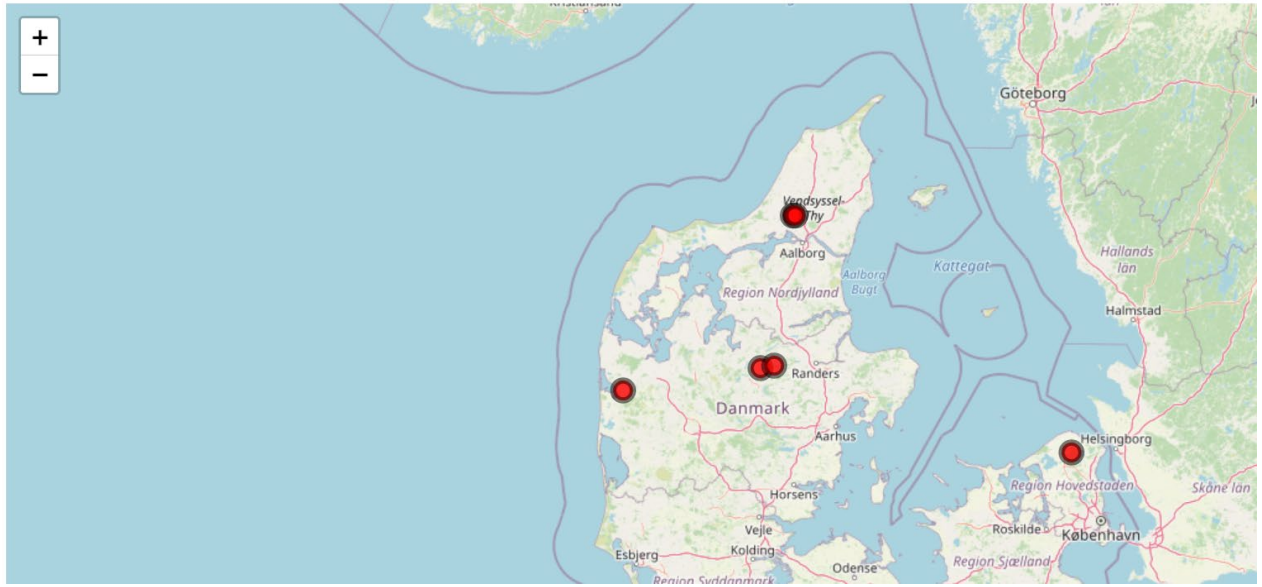
DMYield

6. Platform response after filtering data from the Toolbox



Project Extension – ideas from Ying Wang and Henning Thomsen

Go Into the Data Analysis By Clicking the Map Site



- Can combine with maps to mark the locations of data
- Integrate other projects into the toolbox and shiny R platfrom
- Validate and keep data security by using login user roles
- Benefit that not only Foulumgård staff can input data, but also all relevant researchers who work on the project
- Help to dissemilate the output of the project by using such platform
- Include web access to procedures and protocols, also via mobile phones (Henning)

Personal questions – notes on the small post-it's (5 minutes)

A) What are your needs for better management of your data

Amount of data, modern DBs, Safety, Templates & good examples, Time, Simple and clear procedures, do not make it too complicated, Standards, training and education, help from experts

B) What do you see as opportunities for improving the FAIR management of your data

Need a real DB with query, better overview of data, meta data, FAIR data, better safety system, Update existing platforms for new types of datasets, Better documentation of data, Easier to reuse data, create more structured databases

Group discussion – notes on the large post-it's (10 minutes)

C) Discuss who should do what related to the needs and opportunities identified by the colleagues at your table

Hire one person, Guidelines and templates, a data steward in each section, Better than O drive re safety, quarterly check up. We need solutions for the cases that doesn't fit the template (same data used for different projects - maybe with slight changes. Better to be stored together. Massive data that would be replicated).

Recommendations from the committee to the leadership

Challenges about FAIR data management:

From researcher days, staff express their concerns of handling increasing amount of data, they need modern structured databases for storage, better documentation of data (including meta data), easier to reuse data. Data description and handling must follow international standards and procedures. Data must be stored with clear and transparent safety (better than O-drive), they need templates & good examples. They also realise that this takes time, and hours must be allocated to data management. They express the need for simple and clear procedures and that we do not make it too complicated. Training and education is needed and in the daily work and help from experts is requested.

Recommendations from the committee to the leadership

Rec. 1. Be aware of the FAIR data management policy framework at EU, national, AU, faculty and institute level
Actions: Identify policies and roadmaps at Faculty, AU, national and EU level and identify best practices, standards and procedures relevant for Agro e.g. implementation of ERDA and ERDA-SIF. What does it mean for AGRO. Describe opportunities and challenges. Include one or more persons from Agro in the Open Science Forum at the Faculty level – to feed our needs into this and to capture all relevant info about services from the University to Institutes and sections

Rec. 2. Increase information and training about FAIR data management
Actions: AGRO data management committee propose a strategy for information and training at all levels in the institute. Jens inform section heads about recommendations at the section head meeting. Produce a PPT with info and action plan to be used at all section meetings. Present institute plan and roadmap at next institute meeting. Appoint one Data Steward in each section

Rec. 3. Data management plans and best practice in our institute
Actions: For each Agro database make a digital Data Management Plan based on the EU or KU template. Make a training course for all data stewards to implement the DMPs. DMPs for all major datasets accomplished before end of 2022. Identify best practice. The DMPs will be used for: decisions on how the specific dataset / database should be made FAIR as well as feed into the metadata platform.

Recommendations from the committee to the leadership

Rec. 4. Transition phase - Update the rules and regulations re O-Drive and prepare for the new ERDA and ERDA-SIF repositories

Actions: Jens and Camilla update the guidelines and templates for storage on O-drive, and describe the opportunities of ERDA and ERDA-SIF. Three datasets (Soil texture database and wheat rust database and Mette GIS data) will be used for cases during the implementation of ERDA and ERDA-SIF. Identify additional datasets to be transferred to ERDA and ERDA-SIF, start September 2022. Jens influence the implementation of those repositories via AU Open Science Forum and AU FAIR coordination committee to fulfil the needs of AGRO

Rec. 5. Prioritize and allocate time for data management at the project and institute level

Actions: make rules about the inclusion of a budget for data management in larger projects. Budget time for the data stewards to form smaller network groups across sections to work with FAIR tools and services in practice – and get help from the experts. In a transition phase allocate time for internal project on data management to facilitate the implementation of FAIR data management in the institute

Rec. 6. Dataflow from the field: Field trial planning tool reporting tool from the field. How to get data into the systems efficiently. Automisation. Standard protocols for dataflow and data standards. Need to form a group to work on this

Recommendations from the committee to the leadership about FAIR tools and data

Specific challenges:

1. Equipment and software – data flows in the labs and fields as a part of our infrastructure (computers not on internet and software cannot run on new computers)
2. Human resources, dedicated to FAIR data management. Training of existing staff. At group level – feed into GRUS.
3. Clear descriptions. Which data should be saved and guidelines how, when and by whom.
4. A lot of data in paper form. If those data are relevant then we need dedicated time to digitize the data
5. Consider a time frame – start date. Need to have or nice to have. Get overview of old relevant data (paper, floppy disk digital objects). Start with data related to policy support.
6. Meta data platform describing all our data – manpower needed for this.

AOB

Logbøger i lab: Nu I mapper og tekst. Afrapportering i Excel filer til forskerne. Dennis: word fil på data opsamlings server.

Oplæg næste gang mht. Digitale Lab logbøger.

Næste møde: Info fra Jytte

Data Management Committee 13 June, 2022

Agenda:

1. Approval of agenda
2. Messages:
    a. New member of the committee (Lea Kjærgaard Eriksen)
    b. Activities at Faculty and National level / Jens
    c. Other messages from committee members
3. Follow up on recommendations from last meeting and related to the tasks of the committee/ all
    a. Repositories, - possibilities and roadmap / Jens and Maria
    b. Opportunities to develop a proof of concept ‘Field Trial Planning Tool’ / Jens BK & Jens GH
    c. Use of videos for our procedures and protocols / Jens
    d. Response from Institute leadership on our recommendations from last meeting
    e. How to get “flesh on the bones”, priorities and recommendations
4. E-Doc repository – searchable / Christina and Karin
5. Committee website. What else to put there?
6. AOB
    a. Next meeting 29 August 2022 (new meeting date to be decided upon)

Messages

Informations from Birte Christensen-Dalsgaard (Chairman of DMP strategy at AU)

- Digital transformation – Administration, not research
- ERDA and ERDA SIF test cases postponed from June to September
- First meeting in Open Science Forum and sub-committee, DMP coordination is postponed (24 June decisions will be made)
- Torben Brejnegaard Nielsen (ERDA) and Birte Christensen-Dalsgaard (DM) offer to visit our institute about strategy, AU roadmap and opportunities

Informations from DeiC

KB and KU work on a Dataverse

DeiC

DANISH eINFRASTRUCTURE COOPERATION

International	Generic - Highly recognized in Social Sciences	FA(I)(R)	Publish/Disseminate, Archive, Discover & Re-use, Release, Preserve	Harvard Dataverse	Dataverse is a data repository that is widely used within the Social Sciences. Researchers can login with their institutional credentials via WAYF. Data can be made findable by applying discipline-specific metadata schemes and digital object identifiers (DOIs). Data is made reusable by specifying relevant re-use licenses.

DeiC
DANISH eINFRASTRUCTURE COOPERATION

Sekretariatsfunktion

Kommissorium for arbejdsgruppe vedr. "Politikker, værktøjer og forskerstøtte til FAIR data management" i relation til implementering af den nationale strategi for data management baseret på FAIR principper[1]

Arbejdsgruppen (A) skal udarbejde forslag til handlinger, der adresserer indsatser, der kræver en særlig national koordinering af følgende af strategiens indsatsområder:

1. at få faggrupper på tværs af forskningsinstitutionerne til at definere egen-implementering af FAIR-principper.

4. inden for fagområderne at udarbejde anbefalinger af metoder, som sikrer løbende dokumentation af forskningsdata, i takt med at disse indsamles og genereres, således at alle nødvendige informationer er til stede, når resultater skal gøres tilgængelige.

8. at der etableres samarbejde på tværs af forskningsinstitutionerne om tilvejebringelsen af data stewardship forskerstøttefunktion, hvor dette er gavnligt ud fra eksempelvis et ressourcemæssigt hensyn, f.eks. til større tværgående projekter. Større nationale satsninger koordineres under DeiC.

Arbejdsgruppe A (1) (Politikker):

1.a
inden for fagområderne at udarbejde anbefalinger af **metoder, som sikrer løbende dokumentation af forskningsdata**, i takt med at disse indsamles og genereres, således at alle nødvendige informationer er til stede, når resultater skal gøres tilgængelige.
1.b
at få **faggrupper** på tværs af forskningsinstitutionerne til **at definere egen implementering af FAIR-principper**. Mål: At opbygge et katalog med eksempler på **data management planer** (DMP'er), samt **best practices** for beskrivelse af metadata. Der indsamles gode, konkrete eksempler –"**use cases**" –fra mindst to forskellige faggrupper på implementering af FAIR Principperne. Use cases bør omfatte cases både med og uden håndtering af GDPR-problematik. På baggrund af kataloget, udarbejdes templates/skabeloner på DMP'er, som kan formidles bredt. Derudover udarbejdes et **princip-papir** for best practices, workflows og generelt god implementering af FAIR Principperne.

Handlinger:
- Udarbejdelse af protokol for indsamling af use cases (DMP'er og FAIR implementering)
- Indsamling af **use cases**
- Udarbejdelse af **templates/skabeloner på DMP'er**, der med udgangspunkt i udvalgte fagområder, kan facilitere processen med at dokumentere data. Materialet skal give konkrete anvisninger på hvordan dokumentationen skal udarbejdes, så den lever op til FAIR principperne
- **Princip-papir for best practice**/workflow/implementering generelt af **FAIR Principperne**
- Afholdelse af **workshop(s)** med oplæg fra use cases og evt. af **inviterede eksperter**
- Tilretning af materiale til bidrag til **kommunikationsportal**

Monitorering af fremdrift:
- Løbende **evaluering af use cases** og udarbejdet materiale for relevante faggrupper/potentielle brugere
- Udarbejdelse af lignende løsninger til en bredere gruppe af fagområder vil blive prioriteret på længere sigt

Prioritering af mål og handlinger:
Gode use cases er afgørende for videns-overførsel og smidigere implementering for de forskellige fagområder. Rækkefølgen af de mål og handlinger, beskrevet ovenfor, afspejler dette

Arbejdsgruppe B(6):

- *Juni 2022* Oversigt over interessenter of institutioner med DMP-skabeloner (inkl. kontakt info)
- *30. Juni 2022* Oversigt over DMP-skabeloner fra div institutioner (katalog overeksisterende DMPer).
- *August 2022* Udfærdiger spørgeskema omkring DMPer og FAIR samt forbereder en online workshop til afholdelse i september 2022 (Fokus områder: 1) Diskussion af muligheder for en koordineret indsats ift. den nationale tjeneste med at udarbejde data management skabeloner, 2) indarbejdelse af FAIR i DMP-skabeloner og gøre skabelonerne FAIR i sig selv).
- *September 2022* Afholdelse af workshop (inklusiv interessenter fra Rigsarkivet (evt. Jens Begtrup) og evt. interesserede parter fra Nationalt RDM-netværk).
- *Oktober 2022* Sammenskrivning af dokument med katalog over eksisterende (fagtilpassede) data management skabeloner med forslag til hvordan man indenfor den national tjeneste for data management skabeloner kan udvikle og tilpasse fagtilpassede data management skabeloner, der kan anvendes i forbindelse med anmeldelse til Rigsarkivet.

To do Arbejdsgruppe 1 og B(6)

1. Vi laver 1-3 politikrelaterede spørgsmål som kan kobles på det **spørgeskema** som de planlægger at lave vedr. DMP'er. Disse spørgsmål skal helst være klar indenfor 14 dage.
2. Vi har fokus på at finde **3-4 cases** og når de er fundet bør vi fokusere på hvordan politikker (eller manglende) har understøttet hver enkelt case i at nå i mål med **FAIRificering af data**.
3. Vi laver **workshop** i fællesskab med B(6) – pt. Planlagt til **september.**

PMaaS at AU Tech experimental farming

A vision for an experiment planning tool – planning and asset management

Create a fully customized planning tool

DIS‘ industrial digitalization can be used to create a fully customized experiment planning tool for AU Tech‘s experimental farms in DK

Deeper resolution in experiment planning

The existing system does not allow for planning different sub-experiments on a field. The „lowest“ resolution is a one year cycle of planting and measuring.

Collect manually measured data

Manually collected data and notes related to the experiments can be entered trough simple apps on-site at the field or the tractor.

Automatically collect measurements

Connection field monitors and sensors for automatically collection, storage, preparation and presentation and export.

One task at a time

Focus is on automating the single task that makes the largest positive difference in the workflow. No organisation wide system implementation is needed.

PMaaS at AU Tech - System overview

A vision for an experiment planning tool – planning and asset management

Keep it simple

The point of PMAS is to allow for strategy regarding implementation of a system. In PMAS, the workflow functionality comes first, one task at a time.

Focus is on automating the single task that makes the largest positive difference in the workflow. No organisation wide system implementation is needed

The system infrastructure is prepared for adapting new and diverse features.

Customized to the specific needs

Whatever the scope of the system, the main point of PMAS is to taylor the system to the specific task – or act as a specialized bridge to more rigid systems used by AU Tech.

Supplanting or enhancing exi-sting system

PMAS is designed to work off multiple data sources. Therefore the system can either supplement or replace the existing Nordic Field Trial system – depending on what makes sense.

Multisite from birth

Since the system is based on a Cloud infrastructure, multiple sites are a possibility from day one.

Based on DIS/Microsoft standard infrastructure

A cloud based system from birth. Based on Microsoft Cloud technology, which allow for safe collection and storage of data.

Numerous data treatment services available for analytics of numbers, photos, videos, sounds, etc.

Multiple options for presentation of data and control

What is the first step?

Analysis and prototype

DIS innovative engineering

1 KICK-OFF MEETING

In the kick-off meeting, potential processes are listed and rated. The most relevant ones will be selected for further investigation.

0,5 day

2 PROCESS ANALYSIS

CREADIS performs an analysis of the selected processes, including process walks, interviews, screening of according documents.

1-3 days

3 WORKSHOP

CREADIS presents the findings in a workshop with the relevant stakeholders, along with high-level concepts and expected benefits.

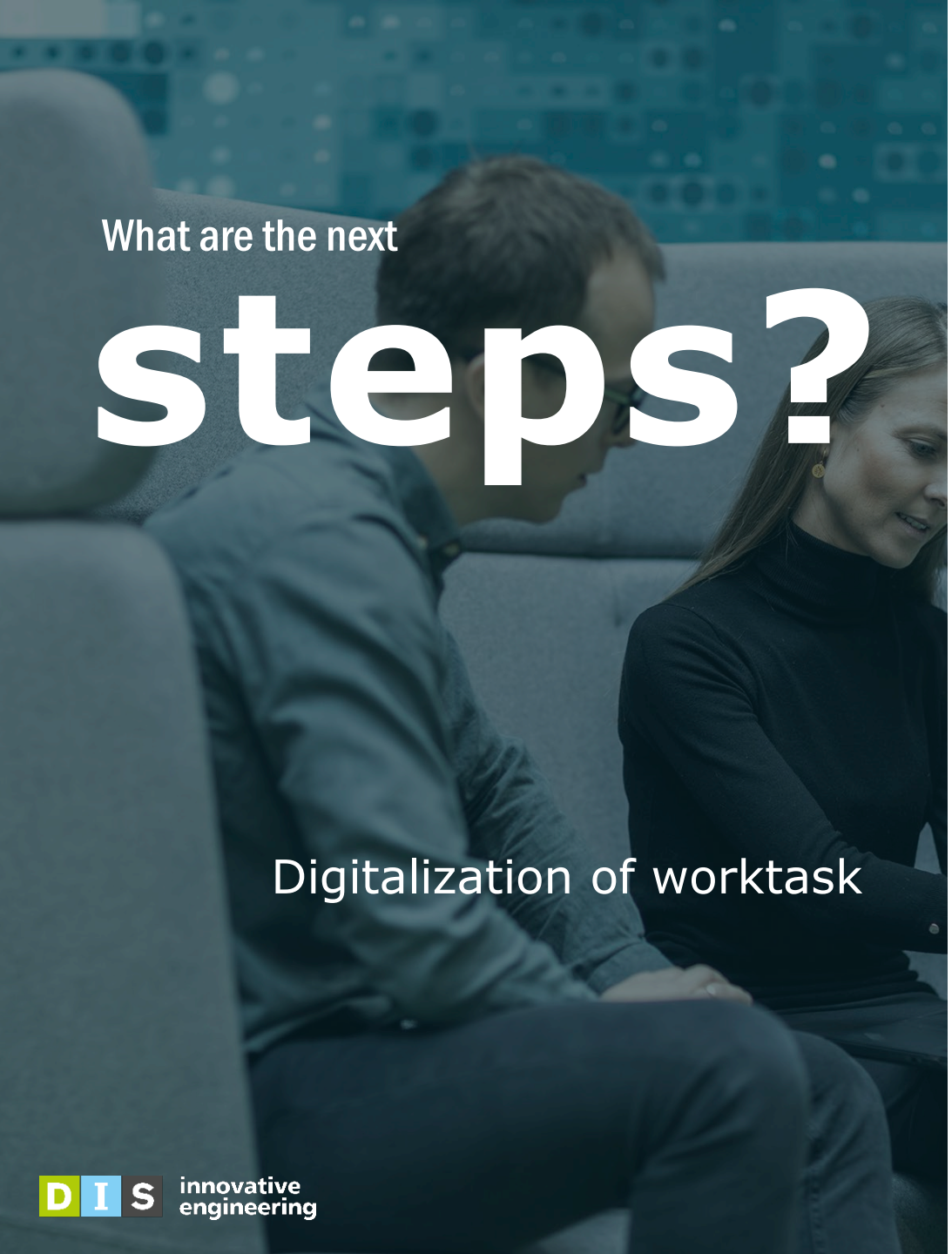
0,5 day

4 MVP / PROTOTYPE

CREADIS builds a working prototype with the sample/core functionality for your processes for your validation. Roadmap and stepwise offers for system development will be presented.

5 days

GET PROOF OF VALUE WITHIN TWO WEEKS



What are the next steps?

Digitalization of worktask

DIS innovative engineering

1 DESIGN & SCOPING

A closer look at a specific worktask, addressing design and scope

2 DEVELOPMENT

Using Scrum methods, development of functionality is performed on all platforms (servers, tablets, field sensors)

3 DEMONSTRATION

DIS presents the system and minor adjustments are agreed and performed

4 ACCEPT AND NEXT TASK

DIS and the system architecture are ready for the next task.

GET PROOF OF VALUE WITHIN TWO WEEKS

Vision

Develop a new workflow and data management system together with DIS, - that Agro can maintain and further improve and develop, alone – but still with DIS as consultants.

Step 1. From field experiments – lab - to data storage, analysis and publication

Define a use case to work on

- Crop Rotation Experiment (As we work on now)
- The Organic crop rotation experiment (Jørgen Eriksens https://dca.au.dk/aktuelt/nyheder/vis/artikel/graes-i-saedskiftet-er-godt-for-klimaet/
- Biobasen (Poul Erik Lærke / Uffe Jørgensen) The Biobasen make use of GIS

ARM is used at Flakkebjerg

Working group members: Anders, Jens Bonderup

Deadline: now

Beskrivelse
KitchenAid 7 cup foodprocessor er rar at have ved hånden, hvad enten du skal hakke, snitte, skære eller purere. Den hjælper dig med det meste og kan håndtere de fleste råvarer. Hårde som bløde, store som små.
Én maskine – mange funktioner
KitchenAid 7 cup foodprocessor kan mere end de fleste køkkenmaskiner. Sæt multiklingen på og brug maskinen til at hakke chokolade, frugter og nødder - enten fint eller groft. Skift til dejkrogen, og ælt dejen til morgendagens boller – nemt og grundigt. Eller sæt rive-/snittejernet på i én af de to størrelser og riv kartofler til röstien eller osten til at drysse på pizzaen.

Kun fantasien sætter grænserne for, hvad foodprocessoren kan hjælpe dig med.
Medfølgende tilbehør
-1,7-liters arbejdsskål (BPA-fri)
-multi-klinge
-dejkrog
-rive-/snittejern til mellemgrov og grov rivning
-nedstopper
-2-i-1-indføringstragt placeret på låget med nedstopper, der har en -smal åbning til madolie
-Drev-adapter til brug af alle klinger og rive-/snittejern

Kom godt i gang
Se video om praktisk brug

Siden sidst

Jordbiologi og Næringsstoffer / Karin Dyrberg

Status:

Laboratoriet:

- Vi har store problemer på laboratoriet pga. af vores apparatur ikke må være på nettet.
- Vi håndtere data på usb-stick fra vores apparatur, og ligger det ind på O- drevet.
- O- drevet er ikke sikkert nok, da alle kan gå ind og rette eller slette .
- Mht til back up på vores apparatur, så gør vi det på en ekstern harddisk, dette er ikke holdbart.
- Pc'er som vores apparatur er koblet op på er meget langsomme pga. at de ikke kan opdateres. Da de ikke er kompatible med nye Windows up dates.

Generelt i gruppen:

Klare guidelines for hvor man gemmer data efter endt PHD forløb og ved færdiggørelse af manuskripter.

2022

M1: DC New server implemented, bought new equipment with a computer without security. Buy computer/server at AU.

M2:

Data management at section CGB / Christina Rønn Ingvardsen

Status:

- At the moment we do not have a database solution for storage. Our main need is a platform to store sequencing data from bigger sequencing projects. We need a platform that can handle row data from sequencing and output from other platforms such as qPCR as well as excel sheets and plain text.
- In general, we have the following type of data:

Sequence data:

- Sequence data from sequencing projects, eg whole genome sequencing, transcriptomics and genotyping. Data from up to approx. 100 plants per experiment.
- This is where we have a need for storage.
- We also have sequence from individual genes from one or more plants, including transgenic plants.

We also have phenotypic data:

- We do not consider most of our phenotypic data as a resource for meta-analysis.
- Phenotypic data might include:
- Expression data (qPCR data) and data on plant quality in general, such as level of resistance to pathogens and insects, protein levels, enzyme activity, plant architecture, ethylene tolerance etc.
- Again, data are from small set of mutants or cultivars (most often) and from a few bigger dataset, approx. 100 plants per experiment.

In CGB we hope for:

a) A clear plan for what should be stored (with a clear triviality limit)
b) A clear plan for who are responsible for storing
c) A storage platform that is easy to use and not least suited for our type of data

2022 M1: Colleagues need more info and help. Problems with PCs and security and access to AU net. Buy AU computers.

2022 M2: Now connected to Workzone. Secure moving of data to Agro folders. Data on O:Drive. Problem when a computer is delivered with lab equipment. Deliver an AU computer to a company and they install the software needed.

Status, needs and possibilities for FAIR management of data in JORD group / Maria Knadel

Data files used:

- Office (Excel, txt, csv), JNB (SigmaPlot), jpeg, R scripts, Access, fasta/fastq
- Geospatial data: shapefiles, raster files in tiff and img
- Images from CT scanner in special file formats and in img
- Spectral data in excel, .nir and OMNIC

Storing data:

- Usually local copies and additional copies on O-drive
- Also stored on: one drive, U-drive, C-drive, H-drive, workstations, personal hard drives
- Sometimes, paper documentation stored for important data sets (finalized projects)

Databases used:

- Access, Excel, txt, ArcGIS Geospatial databases (shapefiles database)

Sharing data with internal and external collaborators:

- Through: O-drive, e-mail, Dropbox, Google drive, WorkZone, ftp, Filesender
- Data of soil databases are not normally shared outside of the department but are free for internal use in the department

Problems/needs related to data management, storage, sharing, archiving, and data flow:

- Sometimes data stored on personal PC only which is not safe if something happens
- Different versions of same files and no metadata available
- GDPR not always considered
- Problems with maintaining databases
- Problems with data handling when PhD students/post docs etc. leave our department. Quality control sometimes lacking
- Sometimes computers in the lab don’t have internet access-need to use a USB drive to save the data
- Access and ArcGIS geodatabases are not user-friendly and problematic as we get more and more data
- The O-drive is a bit ‘messy’ and not easy to navigate, which makes it hard to find data

2022 M1: Plan B re investment on new instruments with computers. Must come with AU computers. How do we measure if we do FAIR data management. Implement Key performance indicators.
M2: How do we measure our success on implementing FAIR. Is our messages going through to the scientists.

Status, needs and possibilities for FAIR management of data in JORD group - continued

Opportunities and ways to improve the status of the above-mentioned issues?

- The Data Management Committee should provide guidelines for metadata generation, how data should be stored and arranged (the template should give only regulations on where data should be stored, but no detailed regulation regarding data format and file system was suggested)
- Regular reminders about GDPR especially in connection with science work would be good
- Access to SAS for statistical analyses should be retained
- Question on security of personal folders on the O drive, a backup in the O drive that is only accessible to the user would be useful
- Suggestion to have a group of technicians across AGRO to work on data management – like in ANIS. Especially in relation to data from automated systems (drones, robots, lab equipment with automatic measurements) or need for allocation of time for database management
- Provide an option to allow external partners access to limited folders – to ease collaboration
- Students not always familiar with data storage/sharing technologies
- Suggestion for a FAIR-management meeting for new Phd-students/employees with an introduction to storing and managing data. The introduction at the GSTS-day is too general and too short, and a bit confusing as there are differences in the procedures in different sections.
- GIS-committee is working on a metadata template to fill out and then place it together with data-maybe we can consider their template.
- When dealing with project data we're often in too much of a hurry and then leave different versions, tests- possible solution a project logbook.

2022 M1:

Patent – Data management. Section rep: Enoch Narh Kudjordjie

Status
We mostly generate and handle the following datasets
•Microbiome sequenced data from Illumina and Oxford nanopore platforms. The Illumina data are mostly in fastq format while oxford nanopore datasets are mostly in fastq and fast5 formats.
•Metabolomics data are saved in excel format
•qPCR data (excel format)
•Plant biomass, environmental parameters, or edaphic data (excel format)
Storage: These datasets are currently stored on the department's O-drive, employee computers or personal drives.
Collaborative research data are sometimes shared with partners

Needs and challenges
We currently need a large, safe, and secure backup data storage system. Storing data on the above-mentioned devices could be compromised. For example, O-drive can be assessed by all employees and data may be mistakenly compromised by other users. Data on personal drives without secured and safe backup storage are at the risk of getting lost.

Possibilities
Data needs to be protected when stored at rest, thus, we gladly welcome the departments common data management platform, which is based on the "FAIR" principles.

2022 M1: Create new folders on O-Drive. Get hold of data from PhD students – what is needed for documentation. Problems with new equipment not on AU network. General problem at AU re security and equipment / computers. Jens ask at Faculty level. Inform the people that help to buy equipment – indkøb (Maria). FTIR – Maria contact Dennis about AU software
2022 M2: Created new folders on O:Drive.

Data management udvalg, Data i Mark Flakkebjerg / Anders Almskou-Dahlgaard

Data fra forsøg i marken:

- I vækstsæson har alle forsøg deres egen markbog i papir, lidt forskellig format efter hvilken gruppe og forsøgstype. Her samles oplysninger der er relevant at have med i marken.
- Data opsamles på forskelligt elektronisk udstyr, Tablets, telefon, bærbar pc og i mindre grad på papir i omtalte markbøger. Ofte på excel-ark på udstyret.
- Det flyttes til netværksdrev, med onedrive, kabeltilslutning eller direkte indtastning.
- Det meste data lægges i ARM (Agricultural Research Manager) da det er programmet, de fleste af de firmaer, vi laver forsøg for, bruger. Data leveres direkte men ofte skal der også laves en printet rapport, hvor forsøgslederen kan sætte sin underskrift på. Kopi af rapporter gemmes på netværksdrev og i printet form i kælderen.
- Alm. Med papirbackup til markbøgererne, som gemmes på arkiv i kælderen. For enkelte gruppe er der op til 60 år gamle små lommebøger på arkivet.
- Analyser af flere tusind korn og frøprøver, gemmes direkte på netværk og i analyseapperatet.
- Der laves talrige tællinger, vejninger og andre analyser af planteprøver, hvor data lagres på forskellig vis; direkte på PC, Tablets og stadig en del på papir, som så skal indtastes bagefter. Både i marken og på laboratorier.

GPS Locations information til markbrug / GIS:

- Der laves faste kørespors information i alle marker. Disse flyttes rundt til alle GPS anlæg, således at alle har mulighed for at ligge deres forsøg korrekt. En person varetager dette og han har data på sin PC/Netværksdrev og flytter info rundt med USB stik.
- Samme person sider med GIS kort over alle marker og kan se markplaner, samt kombineres dette med Dronefoto.

2022 M1:

2022 M2:

Data management udvalg, Data i Mark Flakkebjerg.

Databaser:

- Der findes flere databaser som bruges af medarbejdere i MARK-Flakkebjerg, ofte i samspil med andre grupper som CROP m.fl. og enkelte helt lokale med f.eks. billedinformation. Ikke alle er up to date. Jeg vil beskrive den nok vigtigste for flest i vores gruppe.
- Pest.mdb er en access database der ligger på netværksdrev og som er vital i flere opgaver i Flakkebjerg.
- Historisk set er den startet i 90erne som en FoxPro database over ukrudts forsøgsdata. Den fik jeg konverteret til Access og siden har den udviklet sig også at styrer al kemiregistrering og afvejning i Flakkebjerg.
- Laborant der afvejer kemi til forsøg, registrerer alle indkomne forsøgspræparater.
- Kontorfunktion der modtager alle effektivitets og godkendelsesdokumenter vedrørende kemi, gemmer disse dokumenter på netværksdrev og opretter links i Pest.mdb databasen. Der er mange ældre dokumenter i ringbind på arkiv, som burde skannes ind og linkes ind i DB, for at gører den del historisk komplet.

MARK generelt:

- Alt generelt vedrørende mark, sædskifte og gødningsplaner; føres i Mark Online. Enkelte ting er registreret i Nordic Field trial system.

Diverse:

- Der opbevares mange gamle lysbilleder, desværre uden ret meget info på, vores kælder arkiv – gemmes?
- Forsøgsdata der står på hylderne hos ældre medarbejder der går på pension – Gemmes?
- Forsøgsudstyrsliste/database så man kan se hvem der kan henvises til, når der kommer specielle forespørgsler!

2022 M1:

2022 M2:

Status-needs-possibilities to do better data management – SYSTEM / Mette Vestergaard Odgaard
Types of data in system:
Field/farm experiments (e.g. milk yield), behavioral data (many subjective assumptions), farm data in general (**GDPR issues**), **spatially processed data, landscape data, interviews ! NOT SELF-EXPLANATORY**!

Status: We do not store data on a common (open) drive. Data are located locally or in a messy project folder on O:\ or published in DCA/DCE reports, journals, etc.

Needs: There is a need for a common data structure. Would like easy access to the data of others e.g., the field trials. Still, initially people expressed no need. They go to the person who has the data and information they need or produce data themselves.
But everyone thinks it could be useful with a word/excel-like version of what data people have (also for internally increased collaboration).

Possibilities (and problems): !Data are not self-explanatory + GDPR + founding to do the increased work!

- GDPR. The raw data is often collected with permission from the farmer, it would therefore be time consuming to have to re-collect permits, and some farmers may say now to collaboration because of that.
- It is difficult to clean raw data – self-explainable.
- Time/founding to do better data management.
- If all data should be stored there is a risk of creating a lot of small-matter-data which could be difficult to understand and people probably will not use it → we need a clear ruleset on what kind of data should be stored (only relevant data).
- We should keep respect for each other's competences, and only release data that can be understood – it should be ok to have more detailed data located locally (or what?).
- Maybe produce a script (drejebog) with the FAIR principles that are already being used in the major EU projects – maybe already online (medarbjedere.au).

Example of difficult data: Farm data are currently located on the SQL server where only a few people have access to. Again, here data is not self-explanatory and not suitable for direct further work + GDPR rules often means lower asses to this data + danger that too many users may end up destroying data (do not know if selected locking is possible).
Access to anonymized data on e.g. id15 or 5/10 km grid could be a possibility. Still, this would mean a lot of work. However, Birger, Eva, Christen and Margit are exploring options for a dashboard solutions (similar to John Hopkins covid-19 card).

Overall: Huge willingness to create a metadata-description of existing and future data, to clarify who to go to for more information and to increase interdisciplinary work. Maybe an annual report from each section (responsible person should be the group-representative in data management). This could also be initiated before a larger more complicated setup is in place.

2022 M1: Info and help plus a use case. Information about O-drive and metadata
2022 M2: FAIR principles poster in coffee corner. GDPR sensible data on a SQL server. More people should have access.

Section CROP / Rene Gislum

Status:

Results from grass seed experiments are all in a SAS database. UAV and robot data are on in a folder at the O drive. I am not an IT expert but an IT support helped me and created an folder that apparently are I Roskilde and this has been a great help. I was not able to do image analysis when the images were stored on a server I Aarhus. Back then I have to copy everything fra Flakkebjerg to Aarhus after each flight to secure back-up. Copy it back to Flakkebjerg for image analysis and copy it back to AU for back-up. This problem has been solved now. We furthermore have a lot of images from our VideometerLab (image analysis on seeds) and these are stored locally, which is not ok.

Behov:

Would like to shift to other DB, Data on O-drive not the best. Mirror at AU Roskilde. Data are not safe enough using the O-Drive. Our challenges could be solved by this mirror at Roskilde and this is what I'm currently working on.

Muligheder:

HPC. Data needs to be in Roskilde. All data available for all at the same site for all project members. Talk with ANIS as they have a nice set up for storage of big amount of data.

2022 M1: Cloud computers. Always interact with AU IT.

2022 M2:

Recommendations from the committee to the leadership

Rec. 1. Be aware of the FAIR data management policy framework at EU, national, AU, faculty and institute level
Actions: Identify policies and roadmaps at Faculty, AU, national and EU level and identify best practices, standards and procedures relevant for Agro e.g. implementation of ERDA and ERDA-SIF. What does it mean for AGRO. Describe opportunities and challenges. Include one or more persons from Agro in the Open Science Forum at the Faculty level – to feed our needs into this and to capture all relevant info about services from the University to Institutes and sections

Rec. 2. Increase information and training about FAIR data management
Actions: AGRO data management committee propose a strategy for information and training at all levels in the institute. Jens inform section heads about recommendations at the section head meeting. Produce a PPT with info and action plan to be used at all section meetings. Present institute plan and roadmap at next institute meeting. Appoint one Data Steward in each section

Rec. 3. Data management plans and best practice in our institute
Actions: For each Agro database make a digital Data Management Plan based on the EU or KU template. Make a training course for all data stewards to implement the DMPs. DMPs for all major datasets accomplished before end of 2022. Identify best practice. The DMPs will be used for: decisions on how the specific dataset / database should be made FAIR as well as feed into the metadata platform.

Recommendations from the committee to the leadership

Rec. 4. Transition phase - Update the rules and regulations re O-Drive and prepare for the new ERDA and ERDA-SIF repositories

Actions: Jens and Camilla update the guidelines and templates for storage on O-drive, and describe the opportunities of ERDA and ERDA-SIF. Three datasets (Soil texture database and wheat rust database and Mette GIS data) will be used for cases during the implementation of ERDA and ERDA-SIF. Identify additional datasets to be transferred to ERDA and ERDA-SIF, start September 2022. Jens influence the implementation of those repositories via AU Open Science Forum and AU FAIR coordination committee to fulfil the needs of AGRO

Rec. 5. Prioritize and allocate time for data management at the project and institute level

Actions: make rules about the inclusion of a budget for data management in larger projects. Budget time for the data stewards to form smaller network groups across sections to work with FAIR tools and services in practice – and get help from the experts. In a transition phase allocate time for internal project on data management to facilitate the implementation of FAIR data management in the institute

Rec. 6. Dataflow from the field: Field trial planning tool reporting tool from the field. How to get data into the systems efficiently. Automisation. Standard protocols for dataflow and data standards. Need to form a group to work on this

Recommendations from the committee to the leadership about FAIR tools and data

Specific challenges:

1. Equipment and software – data flows in the labs and fields as a part of our infrastructure (computers not on internet and software cannot run on new computers)
2. Human resources, dedicated to FAIR data management. Training of existing staff. At group level – feed into GRUS.
3. Clear descriptions. Which data should be saved and guidelines how, when and by whom.
4. A lot of data in paper form. If those data are relevant then we need dedicated time to digitize the data
5. Consider a time frame – start date. Need to have or nice to have. Get overview of old relevant data (paper, floppy disk digital objects). Start with data related to policy support.
6. Meta data platform describing all our data – manpower needed for this.

AOB

Logbøger i lab: Nu I mapper og tekst. Afrapportering i Excel filer til forskerne. Dennis: word fil på data opsamlings server. Oplæg næste gang mht. Digitale Lab logbøger.

Discussion point: Agro E-Doc / templates, best practice and good examples

Get flesh on the bones:

- Climate and water developed a proof of concept converting the Crop Rotation Experiment data into a SQL database and including features and tools for effective data flow, data management, display and analysis of the data using an R-platform. Finalyse this and Demonstrate this on next meeting.
- Firefighting: Cents database, Soils databases (GDPR)
- Field trial planning tool including data capture: Proof of concept project with DIS. Start in July.
- Agro E-Doc three use cases, Rene, Anders and Maria
- Preservation and data storage of all material from PhD students / on-boarding and out-boarding. Supervisor should be responsible for data storage and documentation

Milestones:
Start to establish working groups based on a new Internal project
Present our achievements at next Researchers days