Minutes

Jytte Christensen

Date: 7 November 2022
Ref: JYC
WZ: 2021-0293750

———

Page 1/7

AARHUS
UNIVERSITY
DEPARTMENT OF AGROECOLOGY

Meeting date: 24 October 2022, 10:00 – 15:00
Meeting place: Meeting room #3, AU Viborg

Workshop AGRO Data Management Committee

Attendees: Jens Grønbech Hansen (referee for the afternoon program), Anders Almskou-Dahlgaard, Karin Dyrberg, Lea Kjærgaard Eriksen, René Gislum, Christina Rønn Ingvardsen, Maria Knadel, Enoch Narh Kudjordjie, Maarit Mäenpää, Mette Vestergaard Odgaard, Ying Wang, Poul Lassen, Jens Bonderup Kjeldsen, Søren Sommer, Jytte Christensen (referee for the morning program)

AGENDA
1. Welcome and overview (coffee and bread) / Jens
2. Update on AU strategy and FAIR data management including ERDA, ERDA-SIF and DeiC Dataverse / Jens
3. Update from all sections about data management and FAIRification issues. Include a list of all new projects and initiatives where FAIRification is relevant and a potential use case / all Committee members
4. Introduction to the idea of working on project / initiative use cases re FAIRification, Soils databases; BarleyMicroBreed and Crop Disease Platforms / Amelie, Ying, Lea, Poul and Jens
5. Discussion how to work on the use cases and how to include all involved from the field via the lab to the database and reporting / publication / all
6. Collaboration with DIS – what are the needs and ideas / Jens B.K. and Anders
7. Barcoding and best practices from the Lab re. FAIR data management. Experiences and needs / Karin
8. What is next / Jens

Re. 1: Welcome and overview
Jens welcomed all to the first physical meeting in the committee. Lea and Maarit are now members of the committee and they introduced themselves. Ying, Jens B.K., Søren and Poul are special invited for this workshop and they also presented their area of work.

The focus of the workshop today is to discuss how to overcome the barriers for implementation of FAIR data management, identify activities and the next step.

The challenges and needs regarding data management in the department were identified in the spring and recommendations were sent to the department management. The plan was that Jens should attend the latest section leader meeting, but he was prevented from joining. At the AGRO Researcher Days in November, Jens will present the work of the committee and the recommendations as well.

Recommendations to the management is (Short version):

Rec. 1. Be aware of the FAIR data management policy framework at EU, national, AU, faculty and institute level

Actions: Inform about rules and opportunities

Rec. 2. Increase information and training about FAIR data management

Actions: Training plus establish a data management network (Data Stewards)

Rec. 3. Data management plans and best practice in our institute

Actions: Identify Best practice and work on use cases

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

Actions: Start to use ERDA, ERDA-SIF and DeiC Dataverse. Make use the use cases as in 3

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

Actions: Higher budget for Data management in projects. For new projects use the network in 2

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

Actions: Form a working group (Foulum, Flakkebjerg and Askov) to prioritize activities and needs (feed into projects)

Rec. 7. Work on use cases to learn and document the FAIRification process from project development to publication

Actions: Form a network to work on three use cases (learn from each other) and work with ANIS and other groups outside AGRO to start a bottom-up process

Re 2: Update on AU strategy and FAIR data management including ERDA, ERDA-SIF and DEiC Dataverse

Birthe Christensen Dalsgaard is leading the Open Science Forum at AU and a report on implementation of FAIR Data Management at AU is out. Will be added to our website under documents

Under the Open Science Forum, there is a coordination committee on FAIR data management and here is Jens involved. Jens presented the mandate for this committee.
How it all can be organised in practice is not yet clear. At faculty level, the plan is to establish an Open Science Service Centre. The HPC forum, where Maria is member, will be integrated in the service centre infrastructure. More information on Researcher Days.

ERDA and ERDA-SIF should have started in September, but it is postponed to December-January. A meeting on ERDA and ERDA-SIF is planned in Foulum 25 October. (Jens and Lea participate)

AGRO want to be test pilots and Jens would like to have a gross list of up to 10 projects. Jens will start with the DeiC Dataverse (metadata) as soon as possible.

The ERDA and ERDA-SIF data storage platforms is for all project data and we aim at developing data management plans for all projects. Jens will investigate whether external partners can have access to ERDA-SIF.

It is important that budget for data management is set aside in project applications.

It was discussed how to get access to the repositories and how all gain knowledge on this. Data management course could be mandatory for PhD students. It could also be a part of the on-boarding process.

More about this at Researcher days (and this PPT will distributed to all members of our committee and added to our website as well).

Re 3: Update from all sections about data management and FAIRification issues. Include a list of all new projects and initiatives where FAIRification is relevant and a potential use case
The point was discussed as '**Concrete actions for appointing data stewards and creating data management plans**'.

Data stewards:
It was discussed how to appoint and educate data stewards. It was decided to recommend that one data steward at Flakkebjerg and one in Foulum was appointed. Budget must follow these tasks. The two persons could form a dynamic network with relevant persons ad hoc. The data steward must be able to guide and motivate the scientists e.g. on how to navigate in ERDA and ERDA-SIF (a kind of consultancy work). Education of the data steward will be needed.

The responsibility for FAIR data management is on the project leaders.

Data management plans:
We should rely on the AU strategy and try to gain knowledge on making data management plans. Training is needed. As we aim for having data management plans for all projects, this could be incorporated in the cover sheets as a mandatory field.
The aim is to reach a level on FAIR data management expertise where AGRO will become an attractive project partner.

We also have data that does not fit into data management plans. We will accomplish a list of data responsible persons in AGRO.

At our next meeting, Jens will inform about possibilities for education and courses.

Re 4: Introduction to the idea of working on project / initiative use cases re FAIRification, Soils databases; BarleyMicroBreed and Crop Disease Platforms / Amelie, Ying, Lea, Poul and Jens

We will identity use cases, support these and describe the FAIRification process in a way that it is relevant for other projects as well:

#1 Soil databases:
In JORD they have several databases (new and old) and there is a huge amount of data.
The challenges is harmonization and that metadata is either missing or should be translated.
An internal workshop in JORD is held late November to discuss this and to come up with a road-map.
Ying gave a short presentation of the challenges and the work forward (attached)

SQL server capacity was discussed. This should not be a problem as the faculty will offer infrastructure for making data FAIR. The KLIMA section have access to four servers. There might be a fee for using the SQL servers.

Must comply with both the INSPIRE (EU) and the AGRO data management infrastructure.

#2 BarleyMicroBreed
The project is a new EU project and coordinated by AGRO. Lea is involved on the data management part, and it makes good sense to use this project as a use case. There is budget for the work. Jens has been involved as a consultant.

#3 Crop Disease Platforms
Jens gave a short overview of a new Bill Gates funded project to be started in January 2023. Poul presented examples how we in the Toolboxes make data interoperable via WEB Application Programming Interfaces (API) e.g. online calculation of blight risk in Iceland based on historical weather data from a RanchSystems weather station combined with weather forecast data as provided by the Norwegian met office (YR).

Two relevant projects from the SYSTEM section to be added as use cases:
#4 MI BICYCLE: EU project. Mette is appointed as data responsible. She will collaborate with Lea to make data management plan.

#5 MIXED project: Partners have to put data in excel sheets.

From CROP:
#6 Trans4Num: new project from 2023. There will be a lot of satellite data in this project.

It could be nice to have more projects as use cases and describe metadata. Jens would like to have a portfolio of up to 10 projects. He suggested that all committee members asks in their sections about possible projects and reports back to him.

Re 5: Follow up on use cases and more actions needed

Actions

What should be key messages at Researcher days?

How to make your data management plan your user friendly; Status on the DMC Roadmap (Work on use cases, Need to have (EU projects + and nice to have (smaller projects, PhD etc). Part of onboarding. Data Stewards (guidelines and personal help). Jens have made a DMP in the Cover sheet. DMP online (link). Draft presentation by Jens send to all.

Work on Use cases

Three (or more) Use cases – selection, definition - workshop - seminar

How who when: BarleyMicroBreed, Soils databases, Bill gates, MiBiCycle, Trans4Num).

Lea and Jens make a plan.

Jens will contact ANIS about the buttom-up process and next steps at Faculty level

Goal: First datasets at ERDA and ERDA-SIF and at Least 10 datasets or data series described with metadata at DeiC Dataverse. Include also data from RustWatch and Soil spectral library

Re.6: Collaboration with DIS about automated dataflow and trial planning

Jens Bonderup Kjeldsen presented the background (and challenges) for this initiative (see attached). A working group with members from Foulum and Flakkebjerg will come up with recommendation and identified needs

Workshop on describing existing data, needs and recommendation must be accomplished soon

Collaboration with DIS about data capture and flow from field to the lab including Filed Trial planning.

Roadmap

AU Internal documentation of existing databases and data management tools and services
- Data management committee workshop 24 October
- Database workshop second half of November

Status/documentation and identification about the AU needs. Send this to DIS in late November
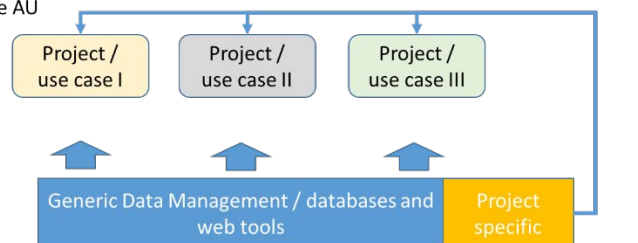
DIS / AU workshop in Mid December
- Discuss AU needs and a system description

Definition of Use Cases (I to III). Online meeting in January and
Report to AU leadership in late January

Realization of our ideas (if funding) in spring 2023

Project / use case I
Project / use case II
Project / use case III

Generic Data Management / databases and web tools
Project specific

Combine generic data management tools with project specific tools

Workshop Agro, AniVet, and CQGG
When we have worked with our use cases we will accomplish a workshop in early spring
between AGRO, ANIVET and QGG and maybe more institutions (whenever relevant).

Re. 7: Barcoding

Need to form a group that can visit TI or other institutes that rely on barcoding all through
the work flow and flow of data. Problem is that some equipment is not able to work with
barcodes. We had a discussion on 'cons' and 'pros' between barcodes and QR codes.

Internal FAIR data management project 2022 and 2023 (Not yet final ap-proved)

Projektet: FAIR management of data and models in AGRO, igangsættes for at ud-
føre konkrete initiativer prioriteret af Agros data management udvalg.

Foreslåede Mål:

1) Draft version af Crop Rotation Experiment er implementeret og dokumenteret i fuld
   skala (Margit, Poul og Ying), jf rapport fra sidste periode.
2) Status, behov og muligheder for flere omlægninger til FAIR data / GDPR sikkerhed: Der skal
   laves en Data Management Plan for alle større projekter/datasæt ved Agro og der skal bru-
   ges relevante skabeloner. Disse DMPere skal identificere hvordan data skal gøres FAIR in-
   klusive lagring, adgang, GDPR følsomhed osv. Der startes med Soils databases (Mogens
   Greve), Integrated grassland-arable crop rotation (Jørgen Eriksen), Biobasen (Uffe og Poul
   Erik), Soil Spectrometry data (Maria Knadel) nyt EU projekt (Lis og Mogens), samt nyt EU
   projekt fra Mogens Nicolaisen (kontakt via Lea). Igangsætning af omlægning for mindst to af
   ovenstående projekter i 2022. Status rapport om det 31/12 2022. Med 2022 erfaring fortsæt-
   tes arbejdet i 2023.
3) System analyse og projektforslag vedr. et Field trial planning, data capture and workflow pro-
   jekt med DIS som konsulent virksomhed. Step 1. Samme med udvalgte Foulumgaard med-
   arbejdere udarbejdes DMPere for alle Access databaser som ønskes omlagt efter FAIR prin-
   cipperne, relevante workflows og protokoller beskrives. Step 2. Der afholdes møde med DIS
   om behov og muligheder for at udvikle et Field trial planning module samt digitalisering og
   effektivisering af data flows fra mark via labs til database inklusive brug af barcodes. Step 3
   DIS og AU udarbejder en systembeskrivelse og Agros ledelse skal tage stilling til go/no go.
   AU skal selv kunne drive og videreudvikle på et sådant system
4) Jens's medlemskab og bidrag til AU Open Science forum / koordineringsudvalg for imple-
   mentering af FAIR data management strategi.
5) Jens's medlemskab og bidrag til arbejdsgruppe vedr. Politikker, værktøjer og forsknings-
   støtte til FAIR data management i relation til implementering af den nationale strategi for
   data management baseret på FAIR principper (her arbejdes også med DMP online og disci-
   plin specifikke skabeloner)
6) Test cases fra Agro vedrørende Implementering af ERDA og ERDA-SIF, udføres december
   2022
7) Test cases fra Agro vedr. DeiC Dataverse (metadata platform) under udvikling i samarbejde
   mellem KU og KB, udføres december 2022
8) Udarbejdelse af Agro strategi for FAIR data management som er i overensstemmelse med
   tilsvarende AU og national strategi, men tilrettet Agros behov inklusive regler, træning og

support for udnyttelse af ERDA, ERDA-SIF og DeiC Dataverse, FAIR er implementeret i onboarding - forløb - og outboarding for PhD og PostDoc. Strategi version 1
Marts 2022