

Referat af DM Rådgivende Forum møde 23.11.2020

Mødedeltagelse:

Medlemmerne af Arbejdsgruppen	Deltager i mødet	Institution
1. Brian Vinter <vinter@au.dk> (Formand)	ja	AU
2. Birte Christensen-Dalsgaard <bcd@cc.au.dk>	ja	AU
3. Karsten Kryger Hansen <kkh@aub.aau.dk>	ja	AAU
4. Diba Terese Markus <dtm@aub.aau.dk>	ja	AAU
5. Claudio Pica <pica@cp3.sdu.dk>	ja	SDU
6. Asger Væring Larsen <avla@bib.sdu.dk>	ja	SDU
7. Mads Sinkær Kjærgaard <mks@ruc.dk>	ja	RUC
8. Sacha Zurcher <szurcher@ruc.dk>	ja	RUC
9. Mareike Buss <mabu.lib@cbs.dk>	ja	CBS
10. Henrik Engell-Hedager <enge@itu.dk>	ja	ITU
11. Hans Henrik Happe <happe@science.ku.dk>	ja	KU
12. Lene Offersgaard <leneo@hum.ku.dk>	ja	KU
13. Nikola Vasiljevic <niva@dtu.dk>	Delvist	DTU
14. Signe Gadegaard <sgad@dtu.dk>	ja	DTU
Sekretariat	Deltager i mødet	Institution
15. Anders Sparre Conrad <anders.conrad@deic.dk>	Ja	DeIC
16. Hannah Mihai <hannah.mihai@deic.dk>	ja	DeIC
17. Rene Belsø <Rene.belso@deic.dk>	Ja	DeIC

1 Velkommst og godkendelse af dagsorden

Formand Brian Vinter bød velkommen og introducerede dagsorden.

Arbejdsgruppen besluttede:

- At godkende dagsorden.

2 Godkendelse af referat fra forrige møde

Signe Gadegaard blev tilføjet som deltager i forrige møde, hvilket fejlagtigt var udeladt.

I referatet fra forrige møde, blev behandlingen af gruppens kommissorium taget op og drøftet indgående.

Formanden hæftede sig ved at DeIC's bestyrelsesformand har udtalt at kommissoriet af gruppen enten kan godkendes eller afvises. Men gruppens opgaver kan derimod justeres, efter indspil fra gruppen selv til bestyrelsen.

Der blev udtrykt forvirring over relevansen af og status på det tidligere DM udredningsarbejde, der allerede var blevet gjort på nationalt niveau (Bilag 04_01 Rapport Data Management i Danmark.pdf, jf dagsordenpunkt 4 nedenfor). Der blev således udtrykt en bekymring for at gruppen over det næste år igen skulle komme med en plan, muligvis gentagelse, for samarbejde mellem universiteterne som er

nationalt og internationalt forankret. Det tidligere omtalte "Internationalt forum" blev også drøftet, men et kommissorium findes fortsat ikke og status er ukendt.

Arbejdsgruppen besluttede:

- At tage drøftelserne til efterretning og godkende referatet.
- At konstatere, at der er behov for at vende tilbage til diskussionen af gruppens kommissorium, formål og opgave.

3 National strategi for data management baseret på FAIR principper

Strategiens hovedtræk blev gentaget fra DeiC's webinar¹. Strategien dækker alt forskningsdata som i varierende grad skal gøres FAIR, og retter sig imod forskningsinstitutioner og forskningsfinansierende fonde. Strategien er struktureret omkring 3 overordnede hovedtemaer:

1. **Kulturændring:**

- ☑ FAIR er udtryk for en forandringsproces i forhold til såvel forståelsen af og udvikling i data management, som i forhold til hvordan forskningens resultater (outputs, så som artikler, data, software mv) meriteres og belønnes.
- ☑ Det skal anerkendes af universiteter og fondene at resultater er mangfoldige, at FAIR data er ét af dem, og at forskergrupper er på forskellige stadier i deres vej til FAIR data.
- ☐ Implementering af en FAIR data praksis skal ske på såvel lokalt, nationalt som internationalt (europæisk/EOSC) niveau.

2. **Teknisk infrastruktur:**

- ☑ Alle forskere skal have adgang til teknisk DM infrastruktur, samt viden om hvor og hvordan forskningsdata og metadata gemmes.
- ☑ Det er behov for såvel basal infrastruktur, finansieret og tilvejebragt af forskningsinstitutionerne, som mere omfattende og/eller kompleks infrastruktur, koordineret og finansieret nationalt.
- ☑ Infrastrukturen skal således hænge sammen lokalt, nationalt og internationalt – teknisk, organisatorisk og økonomisk.

3. **Infrastruktur kompetence:**

- ☑ Solide støttefunktioner – en *vidensinfrastruktur* – er nødvendig, da det ikke er hensigtsmæssigt at forskerne også skal være eksperter i datainfrastruktur og -håndtering.
- ☑ På de højere faglige niveauer er der behov for national koordinering af en kritisk masse af data steward kompetancer.
- ☑ Fondene bør engagere sig i betalingen for vidensinfrastrukturen ved fx data stewards på de højere niveauer.

Strategien er nu i høring bl.a. hos DM Rådgivende Forum's hjemmeinstitutioner, og en ny Styregruppe er planlagt til at udarbejde en implementeringsplan.

Der har været et ønske om at få Rigsarkivet til at bidrage til strategiarbejdet i lyset af Arkivlovens anmeldelsespligt og behovet for at undgå dobbeltarbejde. Rigsarkivet er dog ikke en primær målgruppe.

Arbejdsgruppen besluttede:

- At tage orienteringen til efterretning.
- På næste møde at diskutere begreber og hovedtemaer fra FAIR strategien.

¹ https://www.deic.dk/da/events/webinar-FAIR_datamanagement

4 Præsentation af sidste års *Datamanagement i Danmark* rapport

Birte Christensen-Dalsgaard, AU, præsenterede denne rapport som blev udarbejdet af en arbejdsgruppe som hun var formand for. Rapporten (bilag 04_01) blev, sammen med den tilsvarende rapport² for HPC området, overbragt til DeiC's bestyrelse medio 2019, og kommer med anbefalinger til håndtering af en lokalt, nationalt og internationalt sammenhængende storage-infrastruktur. Rapporten peger bl.a. på mindre forskergruppers behov for en integreret og tidssvarende datahåndteringsinfrastruktur, som meget gerne skulle integrere såvel storage gennem projektperioden, adgang til computing faciliteter og med adgang til brugervenlige tjenester, herunder *trusted repositories*. Rapporten peger også på det ønskelige i at oprette et Back Office funktion, som kan stå for udbud og opfølgning samt til at koordinere samarbejdet på nationalt niveau.

Gruppen fandt at disse konklusioner fortsat var relevante og tilføjede at de skal ses i et internationalt perspektiv.

Arbejdsgruppen besluttede:

- At tage orienteringen til efterretning.

1 DM by Design

Oplæg blev præsenteret af Karsten Kryger Hansen, AAU. DM by Design var et projekt mellem forskningsinstitutioner (finansieret af DM Forum midler), om reproducerbarhed og forudsigelighed. Projektet var om konsensus om begreber og en modellering af aktørers mål og motivation.

Det blev konkluderet at arbejdet er et godt eksempel på at "*one size fits all*" ikke gælder data management infrastruktur.

Arbejdsgruppen besluttede:

- At tage orienteringen til efterretning.

2 Præsentation af eksisterende løsninger

Claudio Pica præsenterede UCloud og Hans Henrik Happe præsenterede ERDA platformen.

UCloud³ er en data centrisk platform for interaktiv HPC, med national adgang gennem WAYF login og fokus på brugervenlighed. UCloud er bygget fra et sæt open source, hvor nye features tilføjes hurtigt, pt. ser man på forskellige roller som administrator, data steward etc. Forskergrupper kan arbejde med og dele data og metadata i sikre projektbaserede virtuelle arbejdspladser. As the software which powers the DeiC Type1 HPC facility, UCloud also supports running applications, such as Jupyter, RStudio, Matlab, Spark, etc. UCloud is covered by the ISO27001 certification of the SDU eScience center.

ERDA⁴ har en fil- og folderstruktur, hvor brugere kan få adgang via mount og web-access til meget store datasæt (op til 1TB). Folders kan deles internationalt via share-links, også med brugere uden ERDA account. Når data skal arkiveres, bliver det gemt på disk eller tape. Da share-links kan mountes, kan ERDA bruges som ren data-backend. Lige nu arbejdes der på integration af JupyterLab og Rstudio. SIF (Sensitive Information Facility) er en (mere) sikker ERDA variant til sensitive persondata. ERDA og SIF bygger på open source POSIX-komponenter.

Der blev spurgt til hvorfor **Sciencedata**⁵ ikke sammen med UCloud og ERDA indgår i gruppens arbejde, hvortil formanden oplyste at DeiC's bestyrelse har besluttet at afvikle Sciencedata, hvorfor det ikke syntes aktuelt at skulle overveje Sciencedata.

² <https://www.deic.dk/da/news/2019-10-16/rapporter>

³ <https://cloud.sdu.dk/app/login>

⁴ <https://www.erda.dk/>

⁵ <https://sciencedata.dk/>

[Efterfølgende præciseret af formanden: Efter indstilling fra CiO forum har DeiCs bestyrelse besluttet ikke at finansiere sciencedata længere. DTU har valgt at videreføre projektet, der leverer ind til et EU projekt. Det vil være betænkeligt at inkludere sciencedata, når hele CiO gruppen har undsagt det].

Arbejdsgruppen besluttede:

- At tage orienteringen til efterretning.

5 Kommercielle cloud udbydere

Med udgangspunkt i at de kommercielle løsninger er en de facto del af forskernes infrastruktur-hverdag, bør det overvejes om brugen af kommercielle løsninger skal forankres i en nationalstrategisk tækning og samarbejde. Der henvises bl.a. til Storbritannien (JISC) der har nationale rammeaftaler med kommercielle leverandører. Der var en drøftelse af Azure, som mange danske universiteter er storkunder hos, og forholdet mellem den "rå infrastruktur" og overliggende produkter og services. Herunder blev det drøftet om det er værd at gå efter at indgå nationale aftaler om sådanne.

Det blev fremført at man med fordel kan se på snitflader, snarere end på nationale løsninger – uanset om de er kommercielle eller offentlige. Gruppen diskuterede kort konsekvenserne af tilvalgs- og fravalgsstrategier, og bemærkede især at aftaler med kommercielle leverandører kræver en exit-strategi, mhp. at undgå vendor lock-in og/eller for at sikre databeskyttelse (ifm. fx Schrems⁶ II-dommen).

GEÁNT har haft workshop-rækken "*Delivering RDM services (optagelser er tilgængelige⁷)*", bl.a. om valget mellem private og offentlige leverandører.

Arbejdsgruppen besluttede:

- At tage orienteringen til efterretning.
- At emnet tages op igen på næste møde, blandt andet mhp. at overveje nationale DM-løsninger af kommerciel art.

6 Gruppens opgaver og eventuelle underudvalg

En hurtig afklaring af gruppens kommissorium og formål/opgaver blev efterspurgt, ligesom der blev udtrykt en bekymring for, at gruppen vil være for længe om at komme med for u-konkrete anbefalinger, med risiko for at samarbejdet og finansiering hertil svækkes til fordel for andre områder fx HPC eller Netværk.

Formanden konstaterede, med afsæt i diskussionerne under dagsorden punkt 2 (Godkendelse af referat fra forrige møde) om kommissoriet og gruppens formål/opgave, at der bør afsættes mindst 30 minutter på næste møde til en grundig og afklarende diskussion.

Der var enighed om at diskussionen må ses i lyset af den nationale FAIR-strategi, og fordelingen mellem lokale front office og nationale back office tiltag.

Udvalgte infrastrukturer udover de der blev præsenteret under dagsorden punkt 6 (UCloud, ERDA, Sciencedata), blev løftet som centrale:

- Der var enighed om at sammenhæng mellem lokale, nationale og internationale infrastruktur løsninger gjorde at særligt en robust og kompatibel AAI-infrastruktur er altafgørende.
- Pain Points blev fremført som vigtige at kende til, og muligvis bruge som ramme for universitetssamarbejdet.
- Finlands arbejde med "*Sensitive Data⁸*" blev fremhævet som en mulig rollemodel

Arbejdsgruppen besluttede:

⁶ <https://www.version2.dk/artikel/nye-eu-anbefalinger-bombe-under-cloud-markedet-1091655>

⁷ <https://wiki.geant.org/display/EV/RDM-Services>

⁸ <https://research.csc.fi/sensitive-data>

- At afsættes mindst 30 minutter på næste møde til en afklarende og konkluderende diskussion af gruppens kommissorium, formål og opgave
- At planlægge en 30 minutters diskussion af AAI på næste møde,
- At forberede et kort indlæg om Pain Points tilgangen, og hvilke Pain Points der findes i de forskellige forskningsinstitutioner.
- At invitere Finland/CSC til at tale om nationale DM infrastrukturer og -services, herunder ifm. sensitive data.

7 Næste møde

Arbejdsgruppen besluttede:

- At næste møde afholdes medio/ultimo januar 2021.

8 Eventuelt

Der blev ikke taget noget op under dette punkt.